A GEOMETRIC STUDY OF WASSERSTEIN SPACES: HADAMARD SPACES

by

Jérôme Bertrand & Benoît Kloeckner

Abstract. — Optimal transport enables one to construct a metric on the set of (sufficiently small at infinity) probability measures on any (not too wild) metric space X. It is called its Wasserstein space $\mathscr{W}_2(X)$ (note that there are variant of this space depending on an exponent, but here we use only the most common, namely the *quadratic cost*).

In this paper we investigate the geometry of $\mathscr{W}_2(X)$ when X has globally non-positive sectional curvature (what we call a *Hadamard space*). Although it is known that –except in the case of the line– $\mathscr{W}_2(X)$ is not non-positively curved, our results show that $\mathscr{W}_2(X)$ have large-scale properties reminiscent of that of X.

First, we are able to define a geodesic boundary in a very similar way than one does for X itself, notably thanks to an *asymptotic formula* that gives good control on the distance between two geodesic rays.

This understanding of the geodesic rays enables us to prove that if X is a visibility space (which for example includes all CAT(-1) spaces, in particular all simply connected manifolds with uniformly negative sectional curvature), then $\mathscr{W}_2(X)$ does not contain any isometrically embedded Euclidean plane: one says that it has rank 1.

We also show that in many cases (including trees and manifolds), if X is negatively curved then the isometry group of $\mathscr{W}_2(X)$ is as trivial as it can be. This contrasts with previous results in the Euclidean case.

While optimal transport has commonly been studied under lower curvature bounds (or to define such bounds), we hope that this paper will show that it is of interest in a non-positively curved setting. Notably, an uncommon feature of our work is that we allow X to be branching, and some emphasis is indeed given to trees.

Contents

1. Introduction 2

Part I. Recalls and notations 5
2. Hadamard spaces 6
3. The Wassertein space 11

Part II. Geodesic rays 16
4. Displacement interpolation for rays 17
5. Asymptotic measures 19
6. Complete geodesics and the rank 22
7. Pairs of boundary measures joined by a geodesic 25

Part III. The geodesic boundary 35
8. Cone topology on $\overline{\mathscr{W}_2(X)}$ 36
9. The boundary of $\mathscr{W}_2(X)$ viewed as a set of measures .. 41

Part IV. Isometric rigidity 44
10. Statement of the result and reduction to a Radon transform 44
11. Injectivity of the Radon transform 51
References 55

1. Introduction

The goal of this paper is to contribute to the understanding of the geometry of Wasserstein spaces. Given a metric space X, the theory of optimal transport (with quadratic cost) gives birth to a new metric space, made of probability measures on X, often called its Wasserstein space and denoted here by $\mathscr{W}_2(X)$ (precise definitions are recalled in the first part of this paper). One can use this theory to study X, for example by defining lower Ricci curvature bounds as in the celebrated works of Lott-Villani [**LV09**] and Sturm [**Stu06**]. Conversely, here we assume some understanding of X and try to use it to study geometric properties of $\mathscr{W}_2(X)$. A similar philosophy underlines the works of Lott in [**Lot08**] and Takatsu and Yokota in [**TY09**].

In a previous paper [**Klo08**], the second author studied the case when X is a Euclidean space. Here we are interested in the far broader class of

Hadamard spaces which are roughly the globally non-positively curved spaces. The first part of the paper is devoted to recalls on optimal transport and Hadamard spaces; in particular the precise hypotheses under which we shall work are given there (Definition 2.1, see also examples 2.2). Let us stress that we allow X to be branching; trees, product involving trees, some buildings are in particular treated in the same framework than, for example, symmetric spaces of non-compact type.

While non-negative curvature is an assumption that is inherited by Wasserstein spaces, it is well-known that non-positive curvature is not (an argument is recalled in Section 3.2). We shall however show that some features of Hadamard spaces still hold in their Wasserstein spaces. Let us now describe the main results of the article.

A Hadamard space admits a well-known geometric compactification, obtained by adding a boundary at infinity made of asymptote classes of geodesic rays. In part II, we study the geodesic rays of $\mathscr{W}_2(X)$. Using a displacement interpolation procedure (Proposition 4.2), we associate to each ray its *asymptotic measure* which lies in a subset $\mathscr{P}_1(c\partial X)$ of probability measures on the cone $c\partial X$ over the boundary of X (Definition 5.1). It encodes the asymptotic distribution of the direction and speed of a measure running along the ray. Our first main result is the *asymptotic formula* (Theorem 5.2) which enables one to compute the asymptotic behavior of the distance between two rays in terms of the Wasserstein distance of the asymptotic measures, with respect to the angular cone distance on $c\partial X$. This asymptotic distance is either bounded or asymptotically linear, so that the boundary $\partial\mathscr{W}_2(X)$ of the Wasserstein space, defined as the set of asymptote classes of unit geodesic rays, inherits an angular metric, just like X does. A striking consequence of the asymptotic formula concerns the rank of $\mathscr{W}_2(X)$, and partially answers a question raised in the previous paper cited above.

Theorem 1.1. — *If X is a visibility space (e.g. if it has curvature bounded from above by a negative constant), then it is not possible to embed the Euclidean plane isometrically in $\mathscr{W}_2(X)$.*

In other words, when X is strongly negatively curved –which implies that its has rank 1–, then although $\mathscr{W}_2(X)$ is not negatively curved it has rank 1 too. Note that our large-scale method in fact implies more general non-embedding results, see Proposition 6.3 and the discussion below.

It is easy to see that any element of $\mathscr{P}_1(c\partial X)$ is the asymptotic measure of some ray in $\mathscr{W}_2(X)$. However, asymptotic measures of complete

geodesic must be concentrated on the set of *unit* geodesics (Proposition 6.2) and this is an easy but important step in the proof of Theorem 1.1. A complete geodesic in $\mathscr{W}_2(X)$ defines two asymptotic measures, also called its ends: one for $t \to -\infty$, the other for $t \to +\infty$. One can therefore ask which pairs of boundary measures are the ends of some complete geodesic. We give in Section 7 a necessary condition that follows easily from the asymptotic formula, and a complete answer when X is a tree (Theorem 7.10). In fact we give two characterizations of realizable pairs of asymptotic measure, one of them involving a optimal transportation problem related to the Gromov product. This criterion might generalize to arbitrary Gromov hyperbolic spaces, but the case of a general Hadamard space may be challenging.

Part III is devoted to the definition of a so-called *cone topology* on $\partial\mathscr{W}_2(X)$ and $\overline{\mathscr{W}_2(X)} = \mathscr{W}_2(X) \cup \partial\mathscr{W}_2(X)$, see Proposition 8.1. Note that the angular metric alluded to above, however useful and meaningful, does not define a satisfactory topology (just as in ∂X, where the angular metric is usually not separable and can even be discrete). The point is that many monotony properties used in the case of Hadamard spaces hold when one restricts to angles based at a Dirac mass. This enables us to carry out the construction of this topology despite the presence of positive curvature. The main result of this part is the following, restated as Theorem 9.2.

Theorem 1.2. — *The asymptotic measure map defined from $\partial\mathscr{W}_2(X)$ to $\mathscr{P}_1(c\partial X)$ is a homeomorphism.*

Note that the two natural topologies on $\partial\mathscr{W}_2(X)$, namely the cone topology and the quotient topology of the topology of uniform convergence on compact sets, coincide. The set $\mathscr{P}_1(c\partial X)$ is simply endowed with the weak topology (where the topology on $c\partial X$ is induced by the cone topology of ∂X).

The possibility to identify $\partial\mathscr{W}_2(X)$ to $\mathscr{P}_1(c\partial X)$ should be thought of as an interversion result, similar to displacement interpolation. The latter says that “a geodesic in the set of measures is a measure on the set of geodesics”, while the former can be roughly restated as “a boundary point of the set of measures is a measure on the (cone over the) set of boundary points”. Note that $\overline{\mathscr{W}_2(X)}$ is not compact; this is quite inevitable since $\mathscr{W}_2(X)$ is not locally compact.

Last, Part IV adresses the isometry group of $\mathscr{W}_2(X)$. In [**Klo08**] it was proved that the isometry group of $\mathscr{W}_2(\mathbb{R}^n)$ is strictly larger than that of $\mathbb{R}^n$ itself, the case $n = 1$ being the most striking: some isometries of $\mathscr{W}_2(\mathbb{R})$ are exotic in the sense that they do not preserve the shape of measures. This property seems pretty uncommon, and we shall show the following rigidity result.

Theorem 1.3. — *If X is a negatively curved Hadamard manifold or a tree not reduced to a line and without leaf, then $\mathscr{W}_2(X)$ is isometrically rigid in the sense that its only isometries those induced by the isometries of X itself.*

Note that in the case of trees, the assumption that X has no leaf could probably be released but comes out naturally in the proof. We are in fact able to treat more space than enclosed in the previous statement, see Theorem 10.1. In particular the $I_{p,q}$ buildings have isometrically rigid Wasserstein spaces.

In the process of proving 1.3, we also get the following result (Corollary 10.5) that seems interesting by itself.

Theorem 1.4. — *Let Y, Z be geodesic Polish spaces and assume that Y is geodesically complete and locally compact. Then $\mathscr{W}_2(Y)$ is isometric to $\mathscr{W}_2(Z)$ if and only if Y is isometric to Z.*

Note that Theorem 1.3 involves the inversion of some kind of Radon transform, that seems to be new (for trees it is in particular different from the horocycle Radon transform) and could be of interest for other problems.

PART I
RECALLS AND NOTATIONS

As its title indicates, this part contains nothing new. We chose to give quite a lot of recalls, so that the reader familiar with non-positively curved spaces can get a crash-course on Wasserstein spaces, and the reader familiar with optimal transport can be introduced to Hadamard spaces.

2. Hadamard spaces

We start with some definitions and properties of metric spaces, especially the non-positively curved ones. Most properties of Hadamard space that we need are proved in [**Bal95**]. Another more extensive reference is [**BH99**].

2.1. Geodesics. — Let us first fix some conventions for any metric space Y (this letter shall be used to design arbitrary spaces, while X shall be reserved to the (Hadamard) space under study).

Given a continuous curve $c : I \to Y$, one defines its length by

$$\ell(c) = \sup_{t_0 < \cdots < t_k} \sum_{i=1}^{k} d(c_{t_{i-1}}, c_{t_i}) \in [0, +\infty]$$

where the supremum is over sequences of arbitrary size of $t_i \in I$. The curve is said to be *rectifiable* if it has finite length.

A *geodesic* in Y is a curve $\gamma : I \to Y$ defined on some interval I, such that there is a constant v that makes the following hold for all times t, t':

$$d(\gamma_t, \gamma_{t'}) = v|t - t'|.$$

In particular, all geodesics are assumed to be *globally* minimizing and to have constant, non necessarily unitary speed. Of course, any geodesic defined over a segment is rectifiable, and has length equal to the distance between its endpoints. When $v = 0$ we say that the geodesic is constant and it will be necessary to consider this case. We denote by $\mathscr{G}^{T,T'}(Y)$ the set of geodesics defined on the interval $[T, T']$. A *geodesic ray* (or *ray*, or *complete ray*) is a geodesic defined on the interval $[0, +\infty)$. A *complete geodesic* is a geodesic defined on $\mathbb{R}$. The set of rays is denoted by $\mathscr{R}(Y)$, the set of unit speed rays by $\mathscr{R}_1(Y)$ and the set of non-constant rays by $\mathscr{R}_{>0}(Y)$. We shall also denote by $\mathscr{G}^{\mathbb{R}}(Y)$ the set of complete geodesics, and by $\mathscr{G}_1^{\mathbb{R}}(Y)$ the set of unit-speed complete geodesics.

2.2. Non-positive curvature. — A metric space is *geodesic* if any pair of points can be linked by a geodesic. A triangle in a geodesic space Y is the data of three points (x, y, z) together with three geodesics parametrized on $[0, 1]$ linking x to y, y to z and z to x. Given a triangle, one defines its *comparison triangle* $(\tilde{x}, \tilde{y}, \tilde{z})$ as any triangle of the Euclidean plane $\mathbb{R}^2$ that has the same side lengths: $d(x, y) = d(\tilde{x}, \tilde{y})$, $d(x, z) = d(\tilde{x}, \tilde{z})$ and $d(y, z) = d(\tilde{y}, \tilde{z})$. The comparison triangle is defined up to congruence.

A triangle with vertices (x, y, z) is said to satisfy the CAT(0) inequality along its $[yz]$ side (parametrized by a geodesic $\gamma \in \mathscr{G}^{0,1}(Y)$) if for all $t \in [0, 1]$, the following inequality holds:

$$d(x, \gamma(t)) \leqslant d(\tilde{x}, (1-t)\tilde{y} + t\tilde{z}) \tag{1}$$

see figure 1. A geodesic space is said to be locally CAT(0) if every point admits a neighborhood where all triangles satisfy the CAT(0) inequality (along all there sides). When Y is a Riemannian manifold, this is equivalent to ask that Y has non-positive sectional curvature. A geodesic space is said to be globally CAT(0) if all its triangles satisfy the CAT(0) inequality. Globally CAT(0) is equivalent to simply connected plus locally CAT(0). We shall simply say CAT(0) for "globally CAT(0)", but this is not a universal convention.

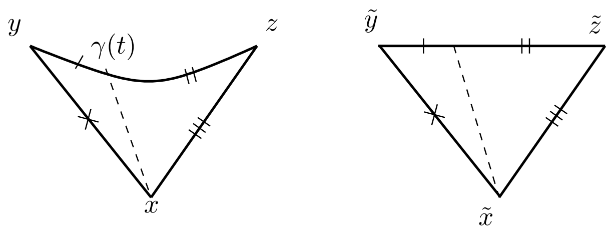

FIGURE 1. The CAT(0) inequality: the dashed segment is shorter in the original triangle on the left than in the comparison triangle on the right.

2.2.1. Angles. — The CAT(0) condition can be translated in terms of angles as follows. Given any geodesic triangle, choose any of its vertices, say x, and assume that the sides containing x are parametrized by two geodesics $\sigma, \gamma \in \mathscr{G}^{0,1}(Y)$. If Y is CAT(0), then the Euclidean angle $\widetilde{\angle}_{\tilde{\gamma}_s\tilde{x}\tilde{\sigma}_t}$ at $\tilde{x}$ is a nondecreasing function of s and t.

One then defines in Y the angle $\angle_{\gamma_1 x \sigma_1}$ at x as the limit, when s and t go to zero, of $\widetilde{\angle}_{\tilde{\gamma}_s\tilde{x}\tilde{\sigma}_t}$. As a consequence, one gets that for any geodesic triangle with vertices (x, y, z) in a CAT(0) space, angle and comparison angle satisfy

$$\angle_{xyz} \leq \widetilde{\angle}_{\tilde{x}\tilde{y}\tilde{z}}.$$

2.2.2. Distance convexity. — In a CAT(0) space, given two geodesics γ and β, the distance function $t \longmapsto d(\gamma_t, \beta_t)$ is convex. This important property shall be kept in mind since it will be used very often in the sequel.

2.3. Hadamard spaces: definition and examples. — We can now introduce the class of spaces we are interested in.

Definition 2.1. — A metric space is a *Hadamard space* if it is:
- Polish (i.e. complete and separable),
- locally compact,
- geodesic,
- CAT(0), implying that it is simply connected.

In all what follows, we consider a Hadamard space X. The Hadamard assumption may not always be made explicit, but the use of the letter X shall always implicitely imply it.

Example 2.2. — There are many important examples of Hadamard spaces. Let us give some of them:
- the Euclidean space $\mathbb{R}^n$,
- the real hyperbolic space $\mathbb{R}\mathrm{H}^n$,
- the other hyperbolic spaces $\mathbb{C}\mathrm{H}^n$, $\mathbb{H}\mathrm{H}^n$, $\mathbb{O}\mathrm{H}^2$,
- more generally the symmetric spaces of non-compact type, like the quotient $\mathrm{SL}(n;\mathbb{R})/\mathrm{SO}(n)$ endowed with the metric induced by the Killing form of $\mathrm{SL}(n;\mathbb{R})$,
- more generally any simply connected Riemannian manifold whose sectional curvature is non-positive,
- trees,
- any product of Hadamard spaces,
- some buildings, like product of trees having unit edges and no leaf or I_{pq} buildings (see [**Bou97, BP99**]),
- the gluing of any two Hadamard spaces along isometric, convex subsets; for example any Hadamard space with an additional geodesic ray glued at some point, or three hyperbolic half-planes glued along their limiting geodesic, etc.

2.4. Geodesic boundary. — The construction of the geodesic boundary that we will shortly describe seems to date back to [**EO73**], but note that [**Bus55**] is at the origin of many related ideas.

Two rays of X are *asymptotic* if they stay at bounded distance when $t \to +\infty$, and this relation is denoted by $\sim$. The asymptote class of a ray γ is often denoted by $\gamma(\infty)$ or γ_∞, and is called the *endpoint* or *boundary point* of γ.

The *geodesic* (or *Hadamard*) *boundary* of X is defined as the set

$$\partial X = \mathscr{R}_1(X)/\sim .$$

Using the convexity of distance along geodesics, one can for example prove that, given points $x \in X$ and $\zeta \in \partial X$, there is a unique unit ray starting at x and ending at ζ.

The union $\bar{X} = X \cup \partial X$ can be endowed with its so-called cone topology, which makes $\bar{X}$ and ∂X compact. Without entering into the details, let us say that this topology induces the original topology on X, and that given a base point x_0 a basic neighborhood of a point $\zeta = \gamma(\infty) \in \partial X$ (where γ starts at x_0) is the union, over all rays σ such that $d(\sigma_t, \gamma_t) < \varepsilon$ for all $t < R$, of $\sigma([R, +\infty])$ (see figure 2).

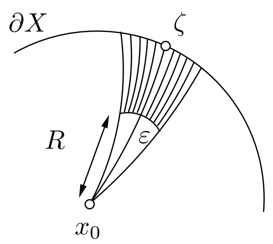

FIGURE 2. A basic neighborhood of a point $\zeta\partial X$, in the cone topology.

Consistently whith the cone topology, all previously defined sets of geodesics, as well as the larger sets $C(I, X)$ of continuous curves defined on an interval I with values in X, are endowed with the topology of uniform convergence on compact sets. With this topology, since X is Hadamard, a geodesic segment is uniquely and continuously defined by its endpoints and a geodesic ray is uniquely and continuously defined by its starting point, its endpoint in the boundary and its speed. As a consequence, there are natural homeomorphisms

$$\mathscr{G}^{T,T'}(X) \simeq X^2, \qquad \mathscr{R}(X) \simeq X \times c\partial X$$

where $c\partial X$ is the cone over ∂X, that is the quotient of $\partial X \times [0, +\infty)$ by the relation $(x, 0) \sim (y, 0)$ for all $x, y \in \partial X$. We usually use the same notation (x, s) for a couple and its equivalence class under this relation; here s shall be considered as a speed. In particular one has $(\mathscr{R}(X)/\sim) \simeq c\partial X$.

Note that in view of our assumptions on X, all these spaces are locally compact Polish topological spaces (that is, metrizable by a metric that

is separable and complete). It ensures that all finite measures on these spaces are Radon.

2.5. Possible additional assumptions. — At some points, we shall make explicitely additional hypotheses on X. One says that a space is:

– *geodesically complete* if every geodesic can be extended to a complete geodesic,
– *non-branching* if two geodesics that coincide on an open set of parameters coincide everywhere on their common definition interval,
– $\mathrm{CAT}(\kappa)$ fo some $\kappa < 0$ if its triangles satisfy the (1) inequality when the comparison triangle is taken in $\mathbb{R}\mathrm{H}^2_\kappa$, the hyperbolic plane of curvature κ, instead of $\mathbb{R}^2$.
– a *visibility space* if for all pairs of distinct points $\alpha, \omega \in \partial X$ there is a complete geodesic γ such that $\gamma(-\infty) = \alpha$ and $\gamma(+\infty) = \omega$.

Note that for all $\kappa < 0$, the condition $\mathrm{CAT}(\kappa)$ implies that X is a visibility space.

Geodesic completeness is quite mild (avoiding examples as trees with leafs), while the other are strong assumptions (for example non-branching rules out trees and visibility rules out products). Another possible assumption, that we shall not use directly, is for X to have *rank one*, meaning that it admits no isometric embedding of the Euclidean plane. It is a weaker condition than visibility. More generally the rank of a space Y is the maximal dimension of an embedded Eulidean space, and it has been proved a very important invariant in the study of symmetric spaces. For example the hyperbolic spaces $\mathbb{R}\mathrm{H}^n$, $\mathbb{C}\mathrm{H}^n$, $\mathbb{H}\mathrm{H}^n$ and $\mathbb{O}\mathrm{H}^2$ are the only rank one symmetric spaces of non-compact type.

2.6. Asymptotic distance. — Given two rays γ, σ, one defines their *asymptotic distance* by

$$d_\infty(\gamma, \beta) = \lim_{t \to +\infty} \frac{d(\gamma_t, \beta_t)}{t}.$$

This limit always exists because of the convexity of the distance function along geodesics. Moreover d_∞ defines a metric on $c\partial X$ and, by restriction, on ∂X (in particular two rays whose distance grows sub-linearly must be asymptotic). It can be proved that d_∞ is the cone metric over ∂X endowed with the angular metric. Namely, for any $(\xi, s), (\xi', t) \in c\partial X$,

$$d_\infty^2((\xi, s), (\xi', t)) = s^2 + t^2 - 2st \cos \angle(\xi, \xi') \tag{2}$$

where $\angle(\xi,\xi') = \sup_{x\in X} \angle_x(\xi,\xi')$ is the supremum of angles between rays issuing from x and asympotic to ξ and ξ' respectively (we refer to [**BH99**, Section II.9] for more details and proof).

It is most important to keep in mind that the metric d_∞ *does not induce the cone topology*, but a much more finer topology. The most extreme case is that of visibility spaces, where $d_\infty(\gamma,\sigma)$ is 0 if $\gamma \sim \sigma$ and the sum of the speeds of γ and σ otherwise: the topology induced on ∂X is discrete. However, the function d_∞ is lower semi-continuous with respect to the cone topology [**BH99**, Proposition II.9.5], so that it is a mesurable function (see the next paragraph for a reference).

In higher rank spaces, it can be useful to turn d_∞ into a length metric, called the Tits metric, but we shall not use it so we refer the interested reader to the books cited above. Let us just note that d_∞ resembles in some aspects the chordal metric on a sphere. In particular, it is naturally isometric to this metric when X is a Euclidean space.

3. The Wassertein space

In this section, we recall the definition of Wassertein space and some of its main properties. For more details, we refer to the books [**Vil03**] and [**Vil09**].

3.1. Optimal transport and Wasserstein metric. — Let us start with the concept of optimal transport which is the theory aimed at studying the *Monge-Kantorovich* problem.

3.1.1. Optimal mass transport problem. — Standard data for this problem are the following. We are given a Polish metric space (Y,d), a lower semicontinuous and nonnegative function $c : Y \times Y \to \mathbb{R}^+$ called the cost function and two Borel probability measures μ,ν defined on Y. A *transport plan* between μ and ν is a measure on $Y \times Y$ whose marginals are μ and ν. One should think of a transport plan as a specification of how the mass in Y, distributed according to μ, is moved so as to be distributed according to ν. We denote by $\Gamma(\mu,\nu)$ the set of transport plans which is never empty (it contains $\mu\otimes\nu$) and most of the time not reduced to one element. The Monge-Kantorovich problem is now

$$\min_{\Pi\in\Gamma(\mu,\nu)} \int_{Y\times Y} c(x,y)\,\Pi(dxdy)$$

where a minimizer is called an *optimal transport plan*. The set of optimal transport plans is denoted by $\Gamma_o(\mu,\sigma)$.

Let us make a few comments on this problem. First, note that under these assumptions, the cost function is measurable (see, for instance, [**Vil03**, p. 26]). Now, existence of minimizers follows readily from the lower semicontinuity of the cost function together with the following compactness result which will be used throughout this paper. We refer to [**Bil99**] for a proof.

***Theorem 3.1* (Prokhorov's Theorem).** — *Given a Polish space Y, a set $P \subset \mathscr{P}(Y)$ is totally bounded (that is, has compact closure) for the weak topology if and only if it is tight, namely for any $\varepsilon > 0$, there exists a compact set K_ε such that $\mu(Y \setminus K_\varepsilon) \leq \varepsilon$ for any $\mu \in P$.*

For example, the set $\Gamma(\mu,\nu)$ is always compact.

We also mention that, compared to the existence problem, the issue of the *uniqueness* of minimizers is considerably harder and requires, in general, additional assumptions. To conclude this introduction, we state a useful criterium to detect optimal transport plan among other plans, named cyclical monotonicity.

***Definition 3.2* (cyclical monotonocity).** — Given a cost function $c : Y \times Y \longrightarrow \mathbb{R}^+$, a set $\Gamma \subset Y \times Y$ is called c-cyclically monotone if for any finite family of pairs $(x_1,y_1),\cdots,(x_m,y_m)$ in Γ, the following inequality holds

$$\sum_{i=1}^{m} c(x_i,y_{i+1}) \geq \sum_{i=1}^{m} c(x_i,y_i) \tag{3}$$

where $y_{m+1} = y_1$.

In other words, a c-cyclically monotone set does not contain cycles of pairs (starting point, ending point) along which a shift in the ending points would reduce the total cost.

***Theorem 3.3*.** — *Let (Y,d) be a Polish space and $c : Y \times Y \longrightarrow \mathbb{R}^+$ be a lower semi-continuous cost function. Then, a transport plan is optimal relatively to c if and only if it is concentrated on a c-cyclically monotone set.*

If c is continuous, this is equivalent to its support being c-cyclically monotone.

Under these assumptions, this result is due to Schachermayer and Teichmann [**ST09**]; see also [**Vil09**] for a proof.

3.1.2. Wassertein space. — Wasserstein spaces arise in a particular variant of the setting above.

***Definition 3.4* (Wassertein space).** — Given a Polish metric space Y, its (quadratic) Wassertein space $\mathscr{W}_2(Y)$ is the set of Borel probability measures μ on Y with finite second moment, that is such that

$$\int_Y d(x_0,x)^2\,\mu(dx) < +\infty \qquad \text{for some, hence all } x_0 \in Y,$$

endowed with the Wasserstein metric defined by

$$\mathrm{W}^2(\mu_0,\mu_1) = \min_{\Pi\in\Gamma(\mu_0,\mu_1)} \int_{Y\times Y} d^2(x,y)\,\Pi(dx,dy).$$

From now on, the cost c will therefore be $c = d^2$.

The fact that W is indeed a metric follows from the so-called "gluing lemma" which enables one to propagate the triangular inequality, see e.g. [**Vil09**].

Remark 3.5. — In this paper, we will also consider the Wassertein space over the cone $c\partial X$ relative to the cost d_∞^2. Since d_∞ is lower semicontinuous, this is indeed a metric space. Note that we shall denote by W_∞ the Wasserstein metric derived from d_∞.

The Wasserstein space has several nice properties: it is Polish; it is compact as soon as Y is, in which case the Wasserstein metric metrizes the weak topology; but if Y is not compact, then $\mathscr{W}_2(Y)$ is not even locally compact and the Wasserstein metric induces a topology stronger than the weak one (more precisely, convergence in Wasserstein distance is equivalent to weak convergence plus convergence of the second moment). A very important property is that $\mathscr{W}_2(Y)$ is geodesic as soon as Y is; let us give some details.

3.2. Displacement interpolation. — The proof of what we explain now can be found for example in chapter 7 of [**Vil09**], see in particular corollary 7.22 and Theorem 7.30. Note that the concept of displacement interpolation has been introduced by McCann in [**McC97**]. We write this section in the case of X, but most of it stays true for any Polish geodesic space.

3.2.1. Dynamical transport plans. — Define a *dynamical transport plan* between two measures $\mu_0, \mu_1 \in \mathscr{W}_2(X)$ as a probability measure μ on $C([t_0, t_1]; X)$ such that for $i = 0, 1$ the law at time t_i of a random curve drawn with law μ is μ_i. In other words we ask $e_{t_i\#}\mu = \mu_i$ where e_t is the map $C([t_0, t_1]; X) \to X$ defined by $e_t(\gamma) = \gamma_t$.

The *cost* of μ is then

$$|\mu|^2 = \int \ell(\gamma)^2 \mu(d\gamma)$$

where $\ell(\gamma)$ is the length of the curve γ (possibly $+\infty$). A dynamical transport plan is *optimal* if it minimizes the cost over all dynamical transport plans, and it is known that a dynamical transport plan exists. Morevover, if μ is an optimal dynamical transport plan then:

(1) the law $(e_{t_0}, e_{t_1})_\#\mu$ of the couple $(\gamma_{t_0}, \gamma_{t_1})$ where γ is a random curve drawn with law μ, is an optimal transport plan between μ_0 and μ_1,
(2) μ-almost all $\gamma \in C([t_0, t_1]; X)$ are geodesics.

Conversely, if Π is a (non-dynamical) optimal transport plan, then one can construct for any $t_0 < t_1$ an optimal dynamical transport plan by the following construction. Let $F : X^2 \to C([t_0, t_1]; X)$ be the map that sends a couple (x, y) of points to the unique geodesic parametrized on $[t_0, t_1]$ that starts at x and ends at y. Then $\mu = F_\#\Pi$ is an optimal dynamical transport plan, whose associated optimal transport plan is obviously Π.

Given a dynamical plan $\mu \in \mathscr{P}(\mathscr{G}^I(X))$ where I is an arbitrary interval, μ is said to be *c-cyclically monotone* if for any $s, t \in I$, the support of the plan $(e_s, e_t)_\#\mu$ is c-cyclically monotone (note that here the cost is continuous). As soon as μ_t have finite second moments, this is equivalent to μ being optimal, but cyclical monotonicity has the advantage of being well-defined without any integrability assumption.

3.2.2. Consequences of optimality. — The main use of optimal dynamical transport plans is that they define geodesic segments. Indeed, let as before e_t be the map $\gamma \mapsto \gamma_t$ defined on the set of continuous curves. If μ is an optimal dynamical transport plan, consider the law $\mu_t = e_{t\#}\mu$ at time t of a random geodesic drawn with law μ: then $(\mu_t)_{t_0 \leqslant t \leqslant t_1}$ is a geodesic of $\mathscr{W}_2(X)$. Displacement interpolation is the converse to this principle.

Proposition 3.6 **(Displacement interpolation)**
Given any geodesic segment $(\mu_t)_{t_0\leqslant t\leqslant t_1}$ *in* $\mathscr{W}_2(X)$*, there is a probability measure* μ *on* $\mathscr{G}^{t_0,t_1}(X)$ *such that for all* t $\mu_t = e_{t\#}\mu$.

If μ is a dynamical transport plan on I, for all $t_0, t_1 \in I$ define the *time restriction* of μ to $[t_0, t_1]$ as $\mu^{t_0,t_1} = r_{t_0,t_1\#}\mu$ where $r_{t_0,t_1}(\gamma)$ is the restriction of the curve γ to the interval $[t_0, t_1]$.

Let μ be an optimal dynamical transport plan on $[0, 1]$. For all $t_0, t_1 \in [0, 1]$ the following hold (see [**Vil09**, Theorem 7.30]):

(1) μ^{t_0,t_1} is an optimal dynamical transport plan,

(2) if X is non-branching and $(t_0, t_1) \neq (0, 1)$, then μ^{t_0,t_1} is the *unique* (up to parametrization) optimal dynamical transport plan between μ_{t_0} and μ_{t_1}.

3.2.3. The Wasserstein space is not non-positively curved. — It is well known that the non-positive curvature assumption on X is not inherited by $\mathscr{W}_2(X)$ except if X is the real line or a subset of it. Let us use displacement interpolation to show it in the most simple way. First, the Wasserstein space is contractible thanks to the affine structure on the set of probablity measures (the convex combination of μ and ν with weights t and $(1-t)$ being $t\mu + (1-t)\nu$). If it were locally CAT(0), it would be globally CAT(0) and, in particular, uniquely geodesic. But as soon as X is not reduced to a geodesic, there exists four distinct points x, x', y, z such that $d(x, y) = d(x, z)$ and $d(x', y) = d(x', z)$ (see figure 3 for the case of a tree). Between the measures $\mu = \frac{1}{2}\delta_x + \frac{1}{2}\delta_{x'}$ (where δ's are Dirac masses) and $\nu = \frac{1}{2}\delta_y + \frac{1}{2}\delta_z$, all transport plans are optimal (and there are several of them, in fact an uncountable number, for example parametrized by the quantity of mass taken from x to y). As we have seen above, each optimal transport plan defines a geodesic in $\mathscr{W}_2(Y)$ from μ to ν. Therefore $\mathscr{W}_2(Y)$ is very far from being uniquely geodesic, and is in particular not non-positively curved.

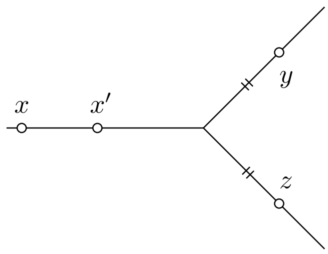

FIGURE 3. Between a measure supported on x and x' and a measure supported on y and z, there are many optimal transport plans and therefore many geodesics.

PART II
GEODESIC RAYS

In this part, we study the geodesic rays of $\mathscr{W}_2(X)$, which shall in particular be used to define its geodesic boundary, to be endowed with a topology in the next part.

The *geodesic boundary* of $\mathscr{W}_2(X)$ is defined as the set of asymptote classes of unitary geodesic rays:

$$\partial \mathscr{W}_2(X) = \mathscr{R}_1(\mathscr{W}_2(X))/\sim .$$

The core result of this part is the asymptotic formula (Theorem 5.2) which gives precise indication on the speed at which the distance between two rays goes to infinity. In particular, we shall show that this distance is either bounded or linear, so that there is a well-defined asymptotic distance in $\partial \mathscr{W}_2(X)$. The asymptotic formula shows that it can be identified with the Wasserstein metric derived from the asymptotic distance of X.

To do this, we need to extend the displacement interpolation principle to rays (Section 4), so that we can define the asymptotic measure of a ray that yields the identification alluded to above (Section 5). Then we use the asymptotic formula to study the rank of $\mathscr{W}_2(X)$ (Section 6), and last we discuss the possibility to prescribe the asymptotic behavior of a complete geodesic for $t \to +\infty$ and $t \to -\infty$ at the same time (Section 7).

4. Displacement interpolation for rays

To start the description of $\partial\mathscr{W}_2(X)$, we need *displacement interpolation* for rays. There is not much, but a little work needed to extend from the case of geodesic segment, the crucial point being to handle the branching case. Note that the case of complete geodesic is not different than that of the rays.

Proposition 4.1 **(Displacement interpolation for rays in the non-branching case)**
If X is non-branching, any geodesic ray (μ_t) of $\mathscr{W}_2(X)$ admits a unique displacement interpolation, that is a probability measure μ on $\mathscr{R}(X)$ such that μ_t is the law of the time t of a random ray having law μ (in other words, such that $\mu_t = (e_t)_\#\mu$).

Note that since (μ_t) is a geodesic, the time restriction of μ to any segment is an optimal dynamical transport plan.

Proof. — Since X is non-branching and (μ_t) is defined for all positive times, we know that there is a unique optimal dynamical transport plan $\tilde{\mu}(T)$ from μ_0 to μ_T, and that $\tilde{\mu}(T) = \tilde{\mu}(T')^{0,T}$ whenever $T < T'$. The fact that X is non-branching also implies that for any two points $x, y \in X$ there is a unique maximal geodesic ray $F(x, y)$ starting at x and passing at y at time 1. Moreover this ray depends continuously and therefore measurably on (x, y). It follows that all $\tilde{\mu}(T)$ are uniquely defined by $\tilde{\mu}(1)$ and that the probability measure $\mu = F_\#(\Pi)$, where Π is the optimal transport plan associated to $\tilde{\mu}(1)$, has the requiered property. □

When X is branching, the previous proof fails for two reasons. The first one is that we cannot determine $\tilde{\mu}(T)$ and define μ from $\tilde{\mu}(1)$ alone; although a theorem of Prokhorov will do the trick. The second problem is that there may exist several optimal dynamical transport plans corresponding to the same geodesic; but the set of these transports is always compact and a diagonal process will solve the problem. However, we lose unicity in the process and it would be interesting to single out one of the dynamical transport plan obtained.

Proposition 4.2 **(Displacement interpolation for rays)**
Any geodesic ray of $\mathscr{W}_2(X)$ admits a displacement interpolation.

Proof. — Let $(\mu_t)_{t\geqslant 0}$ be a geodesic ray in $\mathscr{W}_2(X)$ and for all $T > 0$, let $M(T)$ be the set of all optimal dynamical transport plans parametrized

on $[0,T]$ that induce the geodesic segment $(\mu_t)_{0\leqslant t\leqslant T}$. It is a compact set according to [**Vil09**], Corollary 7.22. For all $T\in\mathbb{N}$, choose $\tilde\mu(T)$ in $M(T)$ and for all integer $0\leqslant T'\leqslant T$ define $\tilde\mu(T'|T)$ as the restriction $\tilde\mu(T)^{0,T'}$.

By a diagonal process, one can extract an increasing sequence T_k of integers such that for all $T'\in\mathbb{N}$, $\tilde\mu(T'|T_k)$ has a limit $\bar\mu(T')\in M(T')$ when $k\to+\infty$. Since for all $T'<T$ the restriction maps

$$p^{T,T'} : \mathscr{G}^{0,T}\to\mathscr{G}^{0,T'}$$

are continuous, we get that $p^{T,T'}(\bar\mu(T))=\bar\mu(T')$. We therefore have a projective system of measures; the projection map

$$p^T : \mathscr{R}(X)\to\mathscr{G}^{0,T}(X)$$

commutes with the $p^{T,T'}$ thus, according to a theorem of Prokhorov (see [**Sch70**]), if we prove tightness, *i.e.* that for all $\varepsilon>0$ there is a compact $K\subset\mathscr{R}(X)$ such that for all T, $\tilde\mu(T)(p^TK)\geqslant 1-\varepsilon$, then we can conclude that there is a unique measure μ on $\mathscr{R}(X)$ such that $p^T_\#\mu=\bar\mu(T)$ for all T. This measure will have the requiered property since $\bar\mu(T)\in M(T)$.

Fix any $\varepsilon>0$. Let K_0,K_1 be compact subsets of X such that $\mu_i(K_i)\geqslant 1-\varepsilon/2$ for $i=0,1$. Let K be the compact subset of $\mathscr{R}(X)$ consisting in all geodesic rays starting in K_0 and whose time 1 is in K_1. Then for all $T>1$, $\tilde\mu(T)(p^T(K))\geqslant 1-\varepsilon$, as needed. □

The following result shall make displacement interpolation particularly useful.

***Lemma 4.3* (lifting)**. — *Let μ, σ be probability measures on $\mathscr{R}(X)$ (or similarly $\mathscr{G}^{T,T'}(X)$, ...) and denote by $\mu_t=(e_t)_\#\mu$ and $\sigma_t=(e_t)_\#\sigma$ their time t.*

Any transport plan $\Pi_t\in\Gamma(\mu_t,\sigma_t)$ admits a lift, *that is a transport plan $\Pi\in\Gamma(\mu,\sigma)$ such that $\Pi_t=(e_t,e_t)_\#\Pi$.*

Proof. — Disintegrate μ along μ_t: there is a family $(\zeta_x)_{x\in X}$ of probability measures on $\mathscr{R}(X)$, each one supported on the set $e_t^{-1}(x)$ of geodesic rays passing at x at time t, such that $\mu=\int\zeta_x\mu_t(dx)$ in the sense that

$$\mu(A)=\int_{\mathscr{R}(X)}\zeta_x(A)\mu_t(dx)$$

for all measurable A. Similarly, write $\sigma=\int\xi_y\sigma_t(dy)$ the disintegration of σ along σ_t.

Define then

$$\Pi(A\times B)=\int_{\mathscr{R}(X)^2}\zeta_x(A)\xi_y(B)\Pi_t(dxdy).$$

It is a probability measure on $\mathscr{R}(X)^2$, and for any measurable sets A,B in X we have $\Pi(e_t^{-1}(A),e_t^{-1}(B))=\Pi_t(A\times B)$ because $\zeta_x\left(e_t^{-1}(A)\right)$ is 1 if $x\in A$, 0 otherwise (and similarly for $\xi_y\left(e_t^{-1}(B)\right)$). A similar computation gives that Π has marginals μ and σ. □

Note that we gave the proof for the sake of completness, but one could simply apply twice the gluing lemma, after noticing that the projection e_t from $\mathscr{R}(X)$ to X gives deterministic transport plans in $\Gamma(\mu,\mu_t)$ and $\Gamma(\sigma,\sigma_t)$.

The lift of Π_t need not be unique; the one constructed in the proof is very peculiar, and can be called the *most independent lift* of Π_t. It is well defined since other disintegration families $(\zeta'_x)_x$ and $(\xi'_y)_y$ must coincide with $(\zeta_x)_x$ and $(\xi_y)_y$ for μ_t-almost all x and σ_t-almost all y respectively.

The lifting lemma shall be used to translate the optimal transport problems between μ_t and σ_t, where these measures move (usually along geodesics) to transport problems between the fixed μ and σ, where it is the cost that moves. In other words, we have just shown that minimizing $\int c(x,y)\,\Pi_t$ over the $\Pi_t\in\Gamma(\mu_t,\sigma_t)$ is the same than minimizing $\int c(\gamma_t,\beta_t)\,\Pi(d\gamma d\beta)$ over the $\Pi\in\Gamma(\mu,\sigma)$.

5. Asymptotic measures

Let us denote by e_∞ the map defined by the formula

$$\begin{array}{rccc} e_\infty: & \mathscr{R}(X) & \longrightarrow & c\partial X \\ & \gamma & \longmapsto & ([\gamma^1],s(\gamma)) \end{array} \tag{4}$$

where γ^1 is the unitary reparametrization of γ, $[\gamma^1]$ is its asymptote class and $s(\gamma)$ is the speed of γ. It is to be understood that whenever $s(\gamma)=0$, $[\gamma^1]$ can be taken arbitrarily in ∂X and this choice does not matter.

Definition 5.1 **(asymptotic measure)**. — Let $(\mu_t)_{t\geq 0}$ be a geodesic ray in $\mathscr{W}_2(X)$ and μ be a displacement interpolation (so that $\mu_t=e_{t\#}\mu$). We define the *asymptotic measure* of the ray by

$$\mu_\infty:=e_{\infty\#}\mu.$$

We denote by $\mathscr{P}_1(c\partial X)$ the set of probability measures ν on $c\partial X$ such that $\int v^2 \nu(dv) = 1$.

In the branching case, the dynamical optimal transport plan is not unique in general. Therefore, the asymptotic measure depends *a priori* on the choice of the dynamical optimal transport plan. We will see soon that it is not the case.

Note that the speed of the geodesic (μ_t) is

$$\left(\int s^2(\gamma)\,\mu(d\gamma)\right)^{1/2} = \left(\int v^2\, s_\# \mu(dv)\right)^{1/2}$$

and we denote it by $s(\mu)$. In particular, $\mathscr{P}_1(c\partial X)$ is the set of measures that correspond to unit speed geodesics. We shall use that the speed function s defined in $\mathscr{R}(X)$ is in $L^2(\mu)$ several times.

The main result of this section is the following.

***Theorem 5.2* (asymptotic formula).** — *Consider two geodesic rays $(\mu_t)_{t\geq 0}$ and $(\sigma_t)_{t\geq 0}$, let μ and σ be any of their displacement interpolations and μ_∞, σ_∞ be the corresponding asymptotic measures. Then (μ_t) and (σ_t) are asymptotic if and only if $\mu_\infty = \sigma_\infty$, and we have*

$$\lim_{t\to\infty} \frac{\mathrm{W}(\mu_t, \sigma_t)}{t} = \mathrm{W}_\infty(\mu_\infty, \sigma_\infty).$$

Therefore: as in X itself, the distance between two rays is either bounded or of linear growth, and two displacement interpolations of the same ray define the same asymptotic measure.

The rôle of the asymptotic formula goes far beyond justifying Definition 5.1 in the branching case: it gives us a very good control on geodesic rays of $\mathscr{W}_2(X)$ on which several of our results rely. To cite one, the asymptotic formula is the main ingredient of Theorem 1.1 on the rank of $\mathscr{W}_2(X)$.

For every $t \geqslant 0$, let d_t be the function defined on $\mathscr{R}(X) \times \mathscr{R}(X)$ by $d_t(\gamma, \beta) = d(\gamma_t, \beta_t)$. We start with an implementation of a classical principle.

***Lemma 5.3*.** — *The function d_t is in $L^2(\Gamma(\mu,\sigma))$, by which we mean that there is a constant $C = C(\mu, \sigma)$ such that for all $\Pi \in \Gamma(\mu, \sigma)$, $\int d_t^2\, \Pi \leqslant C$.*

In the following, it will be of primary importance that C does not depend on Π.

Proof. — Denoting by x any base point in X we have

$$\begin{aligned}\int d_t^2\,\Pi &\leqslant \int \left(d(\gamma_t,x)+d(x,\beta_t)\right)^2\,\Pi(d\gamma d\beta)\\ &\leqslant 2\int d^2(\gamma_t,x)+d^2(x,\beta_t)\,\Pi(d\gamma d\beta)\\ &= 2\int d^2(\gamma_t,x)\,\mu(d\gamma)+2\int d^2(x,\beta_t)\,\sigma(d\beta)\\ &= 2\,\mathrm{W}^2(\mu_t,\delta_x)+2\,\mathrm{W}^2(\sigma_t,\delta_x).\end{aligned}$$

□

Proof of the asymptotic formula. — We shall use several time the following translation of the convexity of the distance function: given any geodesic rays γ,β of X, the function

$$f_t(\gamma,\beta):=\frac{d(\gamma_t,\beta_t)-d(\gamma_0,\beta_0)}{t}$$

is nondecreasing in t and has limit $d_\infty(\gamma,\beta)$.

Assume first that $\mu_\infty=\sigma_\infty$, and let us prove that $\mathrm{W}(\mu_t,\sigma_t)$ is bounded. The lifting lemma gives us a transport plan $\Pi\in\Gamma(\mu,\sigma)$ such that for all (γ,β) in its support, $\gamma_\infty=\beta_\infty$ (simply lift the trivial transport $(\mathrm{Id}\times\mathrm{Id})_\#\mu_\infty$). If $f_t(\gamma,\beta)$ is positive for some t, then $d_\infty(\gamma,\beta)>0$. It follows that on $\operatorname{supp}\Pi$, $d_t\leqslant d_0$. Therefore:

$$\mathrm{W}(\mu_t,\sigma_t)\leqslant\int d_t^2\,\Pi\leqslant\int d_0^2\,\Pi$$

which is bounded by the previous lemma.

Let now Π_∞ be a transport plan from μ_∞ to σ_∞ that is optimal with respect to d_∞. Such a minimizer exists since d_∞ is non-negative and lower-semicontinuous with respect to the cone topology of $c\partial X$; note that taking an almost minimizer would be sufficient anyway. Denote by $\tilde\Pi$ a lift of Π_∞ to $\Gamma(\mu,\sigma)$; then

$$\frac{\mathrm{W}^2(\mu_t,\sigma_t)}{t^2}\leqslant\int\frac{d_t^2}{t^2}\,\tilde\Pi.$$

We have $2\geqslant d_\infty\geqslant f_t\geqslant f_1\geqslant -d_0$ for all $t\geqslant 1$, so that $t^{-1}d_t$ is bounded by $2+d_0$ and $-d_0$. We can thus apply the dominated convergence theorem, which gives

$$\limsup\frac{\mathrm{W}^2(\mu_t,\sigma_t)}{t^2}\leqslant\int d_\infty^2\,\tilde\Pi=\mathrm{W}_\infty^2(\mu_\infty,\sigma_\infty).$$

To prove the other inequality, we introduce $g_t := \max(0, f_t)$. It is a nondecreasing, nonnegative function with d_∞ as limit, and it satisfies $t^2 g_t^2 \leqslant d_t^2$.

Let Π_t be an optimal transport plan between (μ_t) and (σ_t) and $\tilde\Pi_t$ be a lift to $\Gamma(\mu,\sigma)$, which by Prokhorov Theorem is compact in the weak topology. Let $(t_k)_k$ be an increasing sequence such that

$$\lim_k \mathrm{W}(\mu_{t_k}, \sigma_{t_k}) = \liminf_t \mathrm{W}(\mu_t, \sigma_t)$$

and $\tilde\Pi_{t_k}$ weakly converges to some $\tilde\Pi_\infty$.

For all $k' < k$, we have

$$\begin{aligned} \frac{\mathrm{W}^2(\mu_{t_k}, \sigma_{t_k})}{t_k^2} &= \frac{1}{t_k^2} \int d_{t_k}^2 \, \tilde\Pi_{t_k} \\ &\geqslant \int g_{t_k}^2 \, \tilde\Pi_{t_k} \\ &\geqslant \int g_{t_{k'}}^2 \, \tilde\Pi_{t_k}. \end{aligned}$$

Letting $k \to \infty$ and using that the g_t are continuous, we obtain

$$\liminf_t \frac{\mathrm{W}^2(\mu_t, \sigma_t)}{t^2} \geqslant \int g_{t_{k'}}^2 \, \tilde\Pi_\infty$$

for all k'. But $g_{t_{k'}} \leqslant 2$ and the dominated convergence theorem enables us to let $k' \to \infty$:

$$\liminf_t \frac{\mathrm{W}^2(\mu_t, \sigma_t)}{t^2} \geqslant \int d_\infty^2 \, \tilde\Pi_\infty \geqslant d_\infty^2(\mu_\infty, \sigma_\infty).$$

This ends the proof of the asymptotic formula, and shows that if (μ_t) and (σ_t) stay at bounded distance then $\mu_\infty = \sigma_\infty$. $\square$

The end of this part consists in the investigation of two questions where the asymptotic formula finds applications, but from which the sequel is independent. The reader interested in the topology of $\mathscr{W}_2(X)$ can go directly to Part III.

6. Complete geodesics and the rank

In this section we study complete geodesics in $\mathscr{W}_2(X)$, in particular to understand its rank. Recall that the *rank* of a metric space is the highest dimension of a Euclidean space that embeds isometrically in it.

The main result of this section is the following.

Theorem 6.1. — *If X is a visibility space, in particular if it is* $\mathrm{CAT}(\kappa)$ *with $\kappa < 0$, then $\mathscr{W}_2(X)$ has rank* 1.

We expect more generally that for all Hadamard space X, the rank of $\mathscr{W}_2(X)$ is equal to the rank of X, or at least bounded above by some weak rank of X. Theorem 6.1 is a first step in this direction. Note that the fact that we ask the embedding to be nothing weaker than an isometry and to be global is important: it is easy to prove that as soon as X contains a complete geodesic, it admits bi-Lipschitz embedding of Euclidean spaces of arbitrary dimension, and isometric embeddings of Euclidean balls of arbitrary radius and dimension. Both these facts come from an isometric embeeding of open Euclidean half-cone of arbitrary dimension into $\mathscr{W}_2(\mathbb{R})$, see [**Klo08**].

First, let us see that the asymptotic measure of a complete geodesic is much more constraint than that of a mere ray.

Proposition 6.2. — *Let $(\mu_t)_{t\in\mathbb{R}}$ be a complete unit speed geodesic of $\mathscr{W}_2(X)$, and μ be one of its displacement interpolations. Then μ is concentrated on the set of* unit speed *geodesics of X.*

Note that here we do not use any assumption on X (besides the existence of displacement interpolations).

Proof. — Let γ, β be two geodesics in the support of μ and let $a = s(\gamma)$ and $b = s(\beta)$. Fix some point $x \in X$. Then we have the equivalents $d(x, \gamma_t) \sim at$ and $d(x, \beta_t) \sim bt$ when $t \to \pm\infty$. In particular, we get that

$$\begin{aligned} d^2(\gamma_t, \beta_{-t}) &\leqslant \big(d(\gamma_t, x) + d(x, \beta_{-t})\big)^2 \\ &\leqslant (a+b)^2 t^2 + o(t^2) \end{aligned}$$

and similarly $d^2(\beta_t, \gamma_{-t}) \leqslant (a+b)^2 t^2 + o(t^2)$. We also have $d^2(\gamma_t, \gamma_{-t}) = 4a^2t^2$ and $d^2(\beta_t, \beta_{-t}) = 4b^2t^2$. But, since μ is the displacement interpolation of a geodesic, the transport plan Π^t it induces from μ_t to μ_{-t} must respect the cyclical monotonicity. In particular we have

$$d^2(\gamma_t, \gamma_{-t}) + d^2(\beta_t, \beta_{-t}) \leqslant d^2(\gamma_t, \beta_{-t}) + d^2(\beta_t, \gamma_{-t}).$$

From this and letting $t \to \infty$ we get $4a^2 + 4b^2 \leqslant 2a^2 + 2b^2 + 4ab$, which is only possible when $a = b$.

We proved that the speed of geodesics in the support of μ is constant, and since their square integrates (with respect to μ) to 1 we get the desired conclusion. □

As a consequence, the assymptotic measure of a ray that can be extended to a complete geodesic lies in the subset $\mathscr{P}(\partial X)$ of $\mathscr{P}_1(c\partial X)$ (where we identify a space Y with the level $Y \times \{1\}$ in its cone).

Proposition 6.3. — *If X is a visibility space, the space $\mathscr{P}(\partial X)$ endowed with the metric W_∞ (where a ray is identified with its asymptotic measure) contains no non-constant rectifiable curve.*

Proof. — First, the visibility assumption implies that whenever γ and β are non asymptotic, unit speed geodesic rays of X, we have $d_\infty(\gamma,\beta) = 2$. It follows that for all displacement interpolation μ,σ of rays in $\mathscr{W}_2(X)$, assumed to be concentrated on $\mathscr{R}_1(X)$, we have

$$\mathrm{W}_\infty(\mu_\infty,\sigma_\infty) = 2|\mu_\infty - \sigma_\infty|_v^{1/2} \tag{5}$$

where $|\cdot|_v$ is the total variation norm.

The proposition now results from the more general following lemma, which is well-known at least in the case of Euclidean space.

***Lemma 6.4* (Snowflaked metrics)**. — *Let (Y,d) be any metric space and $\alpha < 1$ be a positive number. Then (Y,d^α) is a metric space not containing any non-constant rectifiable curve.*

The fact that d^α is a metric comes from the inequality $a+b \leqslant (a^\alpha + b^\alpha)^{1/\alpha}$ for positive a,b.

Let $c : I \to Y$ be a non-constant curve. Up to restriction and reparametrization, we can assume that $I = [0,1]$ and $c_0 \neq c_1$. Take any positive integer n; since d is continuous, by the intermediate value theorem there are numbers $t_1 = 0 < t_2 < \cdots < t_n < 1$ such that $d(c_{t_{i-1}}, c_{t_i}) = d(c_0,c_1)/n$ and $d(c_{t_n}, c_1) \geqslant d(c_0,c_1)/n$. Denoting by ℓ the length according to the "snowflaked" metric d^α, we get that

$$\ell(c) \geqslant n\left(\frac{d(c_0,c_1)}{n}\right)^\alpha \geqslant d^\alpha(c_0,c_1)n^{1-\alpha}.$$

Since this holds for all n, $\ell(c) = \infty$ and c is not rectifiable. □

Proof of Theorem 6.1. — Assume that there is an isometric embedding $\varphi : \mathbb{R}^2 \to \mathscr{W}_2(X)$. Let r^θ be the ray starting at the origin and making an angle θ with some fixed direction. Then, since r^θ extends to a complete geodesic, so does $\varphi \circ r^\theta$. The displacement interpolation μ^θ of this ray of $\mathscr{W}_2(X)$ must be concentrated on $\mathscr{R}_1(X)$ by Proposition 6.2, so that $\mu^\theta_\infty \in \mathscr{P}(\partial X)$. But φ being isometric, the map $\theta \to \mu^\theta_\infty$ should be an isometric embedding from the boundary of $\mathbb{R}^2$ (that is, the unit circle

endowed with the chordal metric) to $(\mathscr{P}(\partial X), \mathrm{W}_\infty)$. In particular its image would be a non-constant rectifiable curve, in contradiction with Proposition 6.3. $\square$

To prove Theorem 6.1, one could also simply consider the three Euclidean geodesics depicted in figure 4 and check that their relative distances cannot be realized in $\mathscr{P}(\partial X)$ endowed with W_∞. However, our method yields easily more general results: to give two simple examples, Proposition 6.3 rules out the isometric embedding of Minkowski planes ($\mathbb{R}^2$ endowed with any norm), and their cones of the form $\{x^2 < \varepsilon y^2\}$ for any $\varepsilon > 0$. This contrast with above-mentionned fact that even when X is reduced to a line, some Euclidean *half* cones of arbitrary dimension embeds isometrically in $\mathscr{W}_2(X)$.

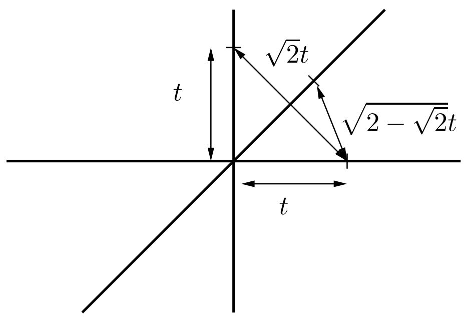

Figure 4. A subset of the plane that cannot appear even asymptotically in $\mathscr{W}_2(X)$.

7. Pairs of boundary measures joined by a geodesic

A complete geodesic (μ_t) in $\mathscr{W}_2(X)$ defines two rays (for $t \leqslant 0$ and $t \geqslant 0$), and one therefore gets two asymptotic measures, denoted by $\mu_{-\infty}$ and $\mu_{+\infty}$, also called the *ends* of the geodesic. By Proposition 6.2, these measures are concentrated on ∂X viewed as a subset of $c\partial X$. In particular, $\mathscr{W}_2(X)$ is already far from being a visibility space. One can ask how far precisely:

Question 1. — *Which pairs of probability measures on ∂X are the ends of some complete geodesic?*

Note that we shall need to consider measures μ on the set of unit complete geodesics $\mathscr{G}_1^{\mathbb{R}}(X)$ that satisfy the cyclical monotonicity, but such that $e_{t\#}\mu$ need not have finite second moment. We still call such maps dynamical transport plan and we say that $e_{\pm\infty\#}\mu$ are its ends. Such a measure defines a complete unit geodesic in $\mathscr{W}_2(X)$ if and only if $e_{t\#}\mu \in \mathscr{W}_2(X)$ for some, hence all $t \in \mathbb{R}$. Note that in this section, we only consider *unit* geodesics even if it is not stated explicitely.

7.1. A first necessary condition: antipodality. — The asymptotic formula shall give us a first necessary condition.

***Definition 7.1* (antipodality).** — Two points $\zeta, \xi \in \partial X$ are said to be *antipodal* if they are linked by a geodesic, that is if $d_\infty(\zeta, \xi) = 2$. Two sets $A_-, A_+ \subset \partial X$ are *antipodal* if all pairs $(\zeta, \xi) \in A_- \times A_+$ are antipodal. Two measures ν_-, ν_+ on ∂X are *antipodal* when they are concentrated on antipodal sets. They are said to be *uniformly antipodal* if their supports are antipodal.

Note that ∂X is compact when endowed with the cone topology, so that the supports of $\nu_\pm$ are compact. If X is a visibility space, then antipodality is equivalent to disjointness; in particular two measures are then uniformly antipodal if and only if they have disjoint supports.

Given a complete unit geodesic μ, the asymptotic formula implies that $\mathrm{W}_\infty(\mu_{-\infty}, \mu_{+\infty}) = 2$, so that *the ends of any complete unit geodesic of $\mathscr{W}_2(X)$ must be antipodal.*

We shall now give a complete answer to Question 1 in the case when X is a tree. In particular, we will see that all pairs of uniformly antipodal measures, but not all pairs of merely antipodal ones, are the ends of a complete geodesic. We shall in fact give two characterizations of ends of complete geodesics, but this will need a fair amount of definitions.

7.2. Flows and antagonism. — From now on, X is assumed to be a tree, described as a graph by a couple (V, E) where: V is the set of vertices; E is the set of edges, each endowed with one or two endpoints in V and a positive length. Since X is assumed to be complete, the edges with only one endpoint are exactly those that have infinite length. Since X is locally compact, as a graph it is locally finite. It is assumed that vertices are incident to 1 or at least 3 edges, so that the combinatorial description of X is uniquely determined by its metric structure. We also fix a base point $x_0 \in X$.

The following definition is the key to characterize cyclically monotone dynamical transport plans.

***Definition 7.2* (antagonism).** — We say that two geodesics are *antagonist* if there are two distinct points x, y such that one of the geodesics goes through x and y in this order, and the other goes through the same points in the other order. This amounts to say that there is an edge followed by both geodesics, but in opposite directions.

The point is that for a dynamical transport plan μ, not having pairs of antagonist geodesic in its support will prove both a very restrictive condition and a characterization of cyclical monotonicity.

We add to each infinite end a formal endpoint at infinity to unify notations. Each edge e has two orientations (xy) and (yx) where x, y are its endpoints. The complement in $\bar{X}$ of the interior of an edge e of endpoints x, y has two components $C_x(xy) \ni x$ and $C_y(xy) \ni y$. An orientation, (xy) say, of an edge has a *future* $(xy)_+ := C_y(xy) \cap \partial X$ and a *past* $(xy)_- := C_x(xy) \cap \partial X$. In other words, the future of (xy) is the set of all possible endpoints of a complete ray passing through x and y in this order.

***Definition 7.3* (Flows).** — Assume ν_- and ν_+ are antipodal measures on ∂X. Define a signed measure by $\nu = \nu_+ - \nu_-$. The *flow* (defined by (ν_-, ν_+)) through an oriented edge (xy) is $\phi(xy) := \nu((xy)_+)$. The flow gives a natural orientation of edges: an oriented edge is *positive* if its flow is positive, *neutral* if its flow is zero, and *negative* otherwise, see figure 5.

Given a vertex x, let $y_1, \dots, y_k$ be the neighbors of x such that (xy_i) is positive, and $z_1, \dots, z_l$ be the neighbors of x such that (xz_j) is negative. Then $\sum_i \phi(xy_i) = \sum_j \phi(z_j x)$ is called the flow through x and is denoted by $\phi(x)$. If $x \neq x_0$, then there is a unique edge starting at x along which the distance to x_0 is decreasing. If this edge is a positive one, (xy_{i_0}) say, then define the *specific flow* through x as $\phi^0(x) = \sum_{i \neq i_0} \phi(xy_i)$. If this edge is a negative one, (xz_{j_0}), then let $\phi^0(x) = \sum_{j \neq j_0} \phi(z_j x)$. If this edge is neutral or if $x = x_0$, then let $\phi^0(x) = \phi(x)$.

Note that since $\{(xy)_+, (xy)_-\}$ is a partition of ∂X and ν has zero total mass, we have

$$\phi(xy) = -\nu((xy)_-) = -\phi(yx).$$

The flow through an edge vanishes if one of the connected component of its complement is finite (but this is not a necessary condition).

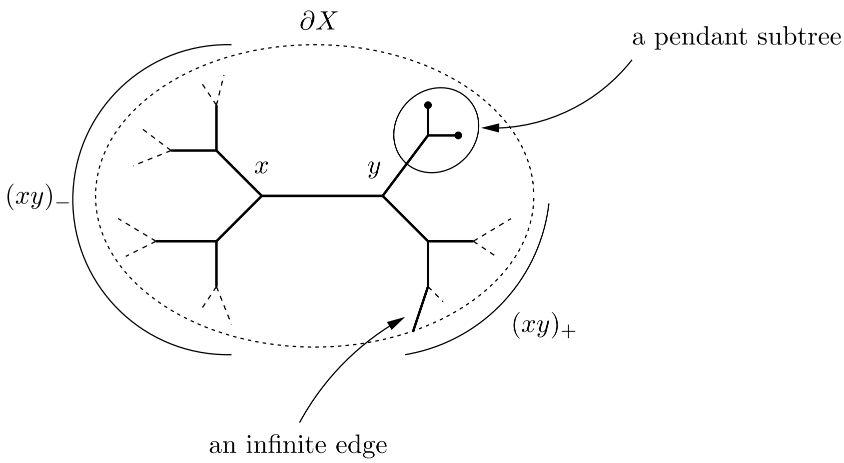

FIGURE 5. The edge (xy) is positive if $\nu_+((xy)_+) > \nu_-((xy)_+)$ since this means that some mass has to move through (xy) (in this direction). The edges of the pendant subtree (finite complement of a vertex) are all neutral.

The next result shows the meaning of flows with regard to our problem. Given a dynamical transport plan μ, we denote by $\mu(xy)$ the μ-measure of the set of geodesics that go through an edge (xy) in this orientation, by $\mu(x)$ the μ-measure of the set of geodesics that pass at x, and by $\mu^0(x)$ the μ-measure of those that are moreover closest to x_0 at this time.

Lemma 7.4. — *If μ is any dynamical transport plan with ends $\nu_\pm$, then:*

(1) *for all edge (xy) we have $\mu(xy) \geqslant \max(\phi(xy), 0)$,*
(2) *for all vertex x we have $\mu(x) \geqslant \phi(x)$.*

and each of these inequality is an equality for all (xy), respectively x, if and only if μ contains no pair of antagonist geodesics in its support. In this case, we moreover have $\mu^0(x) = \phi^0(x)$ for all x.

Note that the set of geodesics that go through an edge (xy) in this orientation is open in $\mathscr{G}^{\mathbb{R}}(X)$, so that μ contains no pair of antagonist geodesics in its support if and only if $\mu \otimes \mu$-almost no pair of geodesics are antagonist.

Proof. — We only prove the first point, the other ones being similar.

Denote by $\mu(C_y(xy))$ the measure of the set of geodesic that lie entrirely in $C_y(xy)$. We have
$$\mu(xy) + \mu(C_y(xy)) = \nu_+((xy)_+) = \phi(xy) + \nu_-((xy)_+)$$
and
$$\nu_-((xy)_+) = \mu(C_y(xy)) + \mu(yx).$$
It follows that $\phi(xy) = \mu(xy) - \mu(yx)$ so that $\mu(xy) \geqslant \phi(xy)$. Moreover the case of equality $\mu(xy) = \max(\phi(xy), 0)$ implies that $\mu(yx) = 0$ whenever $\mu(xy) > 0$, and we get the conclusion. $\square$

From this we deduce in particular a characterization of cyclical monotonicity.

Lemma 7.5. — *A dynamical transport plan μ is d^2-cyclically monotone if and only if $\mu \otimes \mu$-almost no pairs of geodesics are antagonist.*

Proof. — Assume that the support of μ contains two antagonist geodesics γ, β and let x, y be points such that $\gamma_t = x, \gamma_u = y$ where $u > t$ and $\beta_v = y, \beta_w = x$ where $w > v$. Let $r = \min(t, v)$ and $s = \max(u, w)$. Then (see figure 6)
$$d(\gamma_r, \beta_s)^2 + d(\gamma_s, \beta_r)^2 < d(\gamma_r, \gamma_s)^2 + d(\beta_r, \beta_s)^2)$$
so that the transport plan $(e_r, e_s)_\# \mu$ between μ_r and μ_s would not be cyclically monotone.

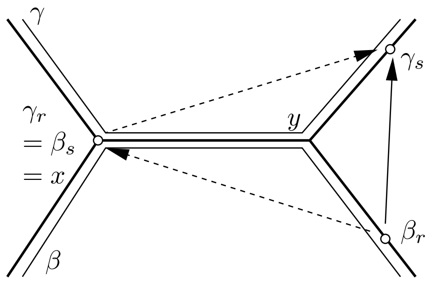

FIGURE 6. Antagonist pairs of geodesics contradict the cyclical monotonicity: the transport plan corresponding to the solid arrow is cheaper than the one corresponding to the dashed arrows.

Assume now that $\mu \otimes \mu$-almost no pairs of geodesics are antagonist. Let $\tau : X \to \mathbb{R}$ be a function that is continuous, increasing isometric on each positive edge and constant on neutral edges. Such a function

can be defined locally around any point, and we can design it globally since X has no cycle. By Lemma 7.4, we see that τ is isometric when restricted to any geodesic in the support of μ (such a geodesic must go through positive edges only). Given times $r < s$, the only $|\cdot|^2$-cyclically monotone transport plan in $\mathbb{R}$ from $\tau_\#\mu_r$ to $\tau_\#\mu_s$ is known to be the increasing rearrangement by convexity of the cost. Here $\tau_\#\mu_s$ is the $r-s$ translate of $\tau_\#\mu_r$, so that this transport plan has cost $(r-s)^2$. But τ is 1-Lipschitz, so that any transport plan from μ_r to μ_s has cost at least $(r-s)^2$. This proves the cyclical monotonicity of μ. □

Note that here for example, μ_r and μ_s need not have finite second moment; however μ induces a transport plan with finite cost between them, and that peculiarity has therefore no incidence on the proof.

7.3. Gromov product. — Before we state the main result of this section, let us turn to a different point of view that might generalize to a wider class of spaces.

Definition 7.6 **(Gromov product)**. — Given $\xi_- \neq \xi_+$ in ∂X, we denote by $(\xi_-, \xi_+) \subset X$ the locus of a geodesic whose ends are ξ_- and ξ_+. Then we write

$$D_0(\xi_-, \xi_+) = d(x_0, (\xi_-, \xi_+))$$

the distance between the base point x_0 and the geodesic (ξ_-, ξ_+). Since X is a tree, this quantity is equal to what is usually called the *Gromov product* $(\xi_- \cdot \xi_+)_{x_0}$, see e.g. [**BH99**]; however the present definition is adapted to our needs. Set $D_0(\xi, \xi) = \infty$ and denote by $\gamma(\xi_-, \xi_+)$ the parametrized unit complete geodesic whose ends are $\xi_\pm$ at $\pm\infty$, and such that its time 0 realizes D_0:

$$d(x_0, \gamma(\xi_-, \xi_+)_0) = D_0(\xi_-, \xi_+).$$

An important property of the Gromov product is that for any $\varepsilon > 0$, $\exp(-\varepsilon D_0)$ metrizes the cone topology on ∂X. In particular, $D_0(\xi_n, \zeta) \to \infty$ if and only if $\xi_n \to \zeta$. Moreover, by compactness of ∂X, if $D_0(\xi_n, \zeta_n) \to \infty$ then there are increasing indices (n_k) such that ξ_{n_k} and ζ_{n_k} converge to a common point. Another obvious consequence is that D_0 is continuous, and one can even check that it is locally constant outside the diagonal. A motivation for using Gromov product here is that it generalizes (with some differences though) to a much more general class of spaces, called Gromov-hyperbolic, including all CAT($\kappa < 0$) Hadamard spaces. One could hope to extend part of Theorem 7.10 below to this class of spaces.

Lemma 7.7. — *The map*

$$\begin{array}{rcl} F : \partial X \times \partial X & \to & \mathscr{G}_1^{\mathbb{R}}(X) \\ (\xi_-, \xi_+) & \mapsto & \gamma(\xi_-, \xi_+) \end{array}$$

is continuous.

Recall that $\mathscr{G}_1^{\mathbb{R}}(X)$ is endowed with the topology of uniform convergence on compact subsets.

Proof. — Since we are dealing with metrizable topologies, we only have to check sequential continuity. Assume $\xi^n \to \xi$ and $\zeta^n \to \zeta$ are two converging sequences in ∂X and set $\gamma = \gamma(\xi, \zeta)$, $\gamma^n = \gamma(\xi^n, \zeta^n)$.

For any $R > 0$, for any large enough n, we have that the rays from γ_0 to ξ^n and to ξ coincide on $[0, R]$. The same property holds true with ζ's. The union of (ξ^n, γ_{-R}), (γ_{-R}, γ_R) and (γ_R, ζ^n) is therefore a geodesic, thus must be the locus of $\gamma(\xi^n, \zeta^n)$. Now $d(\gamma_0, \gamma_0^n)$ is uniformly bounded by $2D_0(\xi, \zeta) + 1$ for n large enough. It follows that γ_0^n lies on γ, so that by definition $\gamma_0 = \gamma_0^n$ and finally γ and γ^n coincide on $[-R, R]$ for n large enough. □

Since F is a right inverse to $(e_{-\infty}, e_{+\infty})$, a transport plan $\Pi \in \Gamma(\nu_-, \nu_+)$ can always be written $(e_{-\infty}, e_{+\infty})_\# \mu$ by taking $\mu = F_\# \Pi$. We shall denote also by D_0 the map $\gamma \mapsto d(x_0, \gamma)$ where γ is any parametrized or unparametrized complete geodesic.

Lemma 7.8. — *A transport plan $\Pi_0 \in \Gamma(\nu_-, \nu_+)$ such that*

$$\int -D_0^2 \, \Pi_0 > -\infty$$

is $-D_0^2$-cyclically monotone if and only if $F_\# \Pi_0$ contains no pair of antagonist geodesics in its support. In this case, Π_0 is a solution to the optimal transport problem

$$\inf_{\Pi \in \Gamma(\nu^-, \nu^+)} \int -D_0^2(\xi, \zeta) \, \Pi(d\xi, d\zeta). \tag{6}$$

Note that the first statement can be recast in the following terms: a dynamical transport plan μ with ends $\nu_\pm$ and such that $\int -D_0^2 \, \mu > -\infty$ is $-D_0^2$-cyclically monotone if and only if it contains no pair of antagonist geodesics in its support.

Proof. — Consider $\Pi_0 \in \Gamma(\nu_-, \nu_+)$ and let $\mu = F_\#\Pi_0$.

Assume first that there are antagonist geodesics γ, β in the support of μ. Then permuting $\gamma_{+\infty}$ and $\beta_{+\infty}$ contradicts the $-D_0^2$-cyclical monotonicity.

To prove the other implication, assume that $\operatorname{supp}\mu$ contains no pair of antagonist geodesics but that Π_0 does not achieve the infimum in (6). This happens notably when Π_0 is not cyclically monotone.

Then there is some $\Pi_1 \in \Gamma(\nu_-, \nu_+)$ such that

$$\int -D_0^2\,\Pi_0 > \int -D_0^2\,\Pi_1.$$

If $F_\#\Pi_1$ has couples of antagonist geodesics in its support, then we can still improve Π_1. Choosing any numbering $e_2, e_3, \dots$ of the non-oriented edges of X, we inductively construct transport plans $\Pi_2, \Pi_3, \dots$ in $\Gamma(\nu_-, \nu_+)$ such that $F_\#\Pi_k$ has no antagonist geodesics through the edges $e_2, \dots, e_k$ in its support, and $(-D_0^2)_\#\Pi_k$ is stochastically dominated by $(-D_0^2)_\#\Pi_{k-1}$ (for all (ζ, ξ) going through e_k in the negative direction, replace ξ by some ξ' in the future of e_k and corresponding to a $(\zeta', \xi') \in \operatorname{supp}\Pi_{k-1}$, and replace (ζ', ξ') by (ζ', ξ); there are many choices to do but they can be made in a arbitrary manner). Then we get

$$\int -D_0^2\,\Pi_{k-1} \geqslant \int -D_0^2\,\Pi_k$$

where some, or even all of this integrals can be negative infinite.

We shall use a weak convergence, but $-D_0^2$ is not bounded; we therefore introduce the functions $f_T = -\min(D_0^2, T)$ for all $T \in \mathbb{N}$. The transport plans Π_k also satisfy

$$\int f_T\,\Pi_{k-1} \geqslant \int f_T\,\Pi_k$$

for all T.

Since ∂X is compact, so is $\Gamma(\nu_-, \nu_+)$ and we can extract a subsequence of (Π_k) that weakly converges to some $\tilde\Pi$, and $\tilde\mu := F_\#\tilde\Pi$ has no pair of antagonist geodesics in its support.

The montone convergence theorem implies that for T large enough, we have $\int f_T\,\Pi_1 < \int -D_0^2\,\Pi_0$, and by weak convergence we get $\int f_T\,\tilde\Pi \leqslant \int f_T\,\Pi_1$. Since $-D_0^2 \leqslant f_T$, we get

$$\int -D_0^2\,\tilde\Pi < \int -D_0^2\,\Pi.$$

But then by Lemma 7.4 we get that $\tilde{\mu}^0(x) = \mu^0(x)$ for all $x \in V$, and since $\int -D_0^2\,\Pi_0 = \sum_x -d^2(x_0,x)\mu^0(x)$ it follows that $\int -D_0^2\,\Pi_0$ and $\int -D_0^2\,\tilde{\Pi}$ must be equal, a contradiction. ☐

Note that the lemma stays true if $-D_0^2$ is replaced with any decreasing function of D_0, but that we shall need precisely $-D_0^2$ later.

Remark 7.9. — In this proof we cannot use Theorem 4.1 of [**Vil09**] since we do not have the suitable lower bounds on the cost.

7.4. Characterization of ends. — We can now state and prove the answer to Question 1 in the case of trees.

Theorem 7.10. — *Assume that X is a tree and let ν_-, ν_+ be two antagonist measures on ∂X. The following are equivalent:*

- *there is a complete geodesic in $\mathscr{W}_2(X)$ with $\nu_\pm$ as ends;*
- *the optimal transport problem* (6) *is finite:*

$$\inf_{\Pi\in\Gamma(\nu^-,\nu^+)} \int -D_0^2(\xi,\zeta)\,\Pi(d\xi,d\zeta) > -\infty;$$

- *the specific flow defined by $\nu_\pm$ satisfies*

$$\sum_{x\in V} \phi^0(x) d(x,x_0)^2 < +\infty.$$

When these conditions are satisfied, then the optimal transport problem (6) *has a minimizer Π_0 and $\Gamma_\# \Pi_0$ define a geodesic of $\mathscr{W}_2(X)$ with the prescribed ends.*

Moreover, the above conditions are satisfied as soon as $\nu_\pm$ are uniformly antipodal.

Proof. — First assume that there is a complete geodesic in $\mathscr{W}_2(X)$ with $\nu_\pm$ as ends and denote by μ one of its displacement interpolations; by Lemma 7.5, the support of μ does not contain any pair of antagonist geodesics. From Lemma 7.4 it follows that

$$\sum_{x\in V} \phi^0(x) d(x,x_0)^2 \leqslant \int_X d(x,x_0)^2 \mu_0(dx) < \infty$$

since by hypothesis $\mu_0 \in \mathscr{W}_2(X)$. We have also

$$\int -D_0^2(\xi,\zeta)\,\Pi_0(d\xi,d\zeta) \geqslant -\int_X d(x,x_0)^2 \mu_0(dx) > -\infty$$

so that Lemma 7.8 implies that $\Pi_0 := (e_{-\infty}, e_{+\infty})_\# \mu$ is a solution to problem 6.

Now consider the case when $\nu_\pm$ are uniformly antipodal. Since the supports of ν^- and ν^+ are disjoint, the map D_0, when restricted to $\operatorname{supp}\nu^- \times \operatorname{supp}\nu^+$, is bounded. Therefore, since it is a continuous map, the optimal mass transport problem is well-posed and admits minimisers.

More generally when the infimum in problem (6) is finite, by using the regularity of Borel probability measures on ∂X we can approximate ν^- and ν^+ by probability measures whose supports are disjoint sets. Then, the previous paragraph gives us a sequence of plans which are $-D_0^2$-cyclically monotone. Since D_0 is a continuous map, Prokhorov's theorem allows us to extract a converging subsequence whose limit Π_0 is $-D_0^2$-cyclically monotone. By the finitness assumption,

$$\int -D_0^2(\xi,\zeta)\,\Pi_0(d\xi,d\zeta) > -\infty$$

and Π_0 is a $-D_0^2$-optimal transport plan.

As soon as a minimizer Π_0 to (6) exists,by Lemma 7.8 $\mu := \Gamma_\# \Pi_0$ is a dynamical transport plan that has no antagonist pair of geodesics in its support. By Lemma 7.5 μ is cyclically monotone. By its definition

$$\int_X d(x,x_0)^2 \mu_0(dx) = \int D_0^2(\xi,\zeta)\,\Pi_0(d\xi,d\zeta) < +\infty$$

so that μ defines a geodesic of $\mathscr{W}_2(X)$. It has the prescribed ends since $\Pi_0 \in \Gamma(\nu_-, \nu_+)$.

We have only left to consider the case when $\sum_{x\in V} \phi^0(x) d(x,x_0)^2 < \infty$. For this, let us construct a suitable complete geodesic by hand. Let τ be the time function such that $\tau(x_0) = 0$, as in the proof of Lemma 7.5. The levels of the time function are finite unions of isolated points and of subtrees of X all of whose edges are neutral. Indeed, consider a point a: if it lies inside a neutral edge, then all the edge has time $\tau(a)$. Otherwise, let (xy) be the orientation of this edge that is positive: points on $[x,a)$ have time lesser than $\tau(a)$, while points on $(a,y]$ have time greater than $\tau(a)$. If a is a vertex, then similarly one sees that nearby a, only the points lying on an incident neutral edge can have time equal to $\tau(a)$. Let $\dot\tau^{-1}(t)$ be the union of the isolated points of the level $\tau^{-1}(t)$ and of the points that lie (in X) on the boundary of the neutral subtrees of the same level. In other words, $\dot\tau^{-1}(t)$ is the level t of the map induced by τ on the subforest of X where all neutral edges have been removed.

Define now
$$\tilde{\mu}_t = \sum_{a\in\dot{\tau}^{-1}(t)} \phi(a)\,\delta_a$$
where $\phi(a) = \phi(xy)$ if a lies inside a positive edge (xy). Note that $\tilde{\mu}_t$ is a probability measure thanks to the antipodality of ν_- and ν_+: without it, it would have mass less than 1. It is a good first candidate to be the geodesic we are looking for, except the second moment of μ_t need not be finite! To remedy this problem, proceed as follows. First, there is a displacement interpolation $\tilde{\mu}$ of $(\tilde{\mu}_t)$, which is a probability measure on $\mathscr{G}_1^{\mathbb{R}}(X)$. Now, construct a random geodesic γ as follows : draw $\tilde{\gamma}$ with law $\tilde{\mu}$, and let γ be the geodesic that has the same geometric locus and the same orientation as $\tilde{\gamma}$, and such that γ is nearest to x_0 at time 0. The condition $\sum_{x\in V}\phi^0(x)d(x,x_0)^2 < \infty$ ensures that the law of γ_0 has finite second moment. □

The example shown in Figure 7 shows that antipodality is not sufficient for $\nu_\pm$ to be the ends of a geodesic.

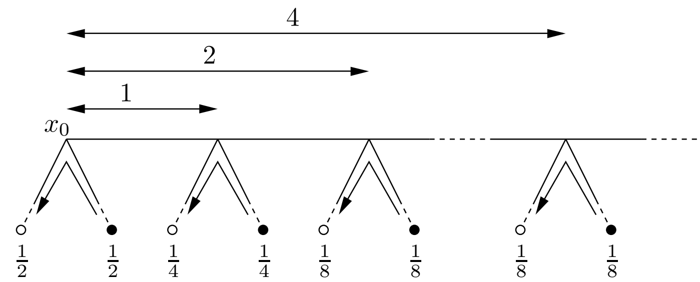

FIGURE 7. The measures ν_- (black dots) and ν_+ (white dots) are antipodal. However, the only possible geodesics having these measures as ends, depicted by simple arrows, are not in $\mathscr{W}_2(X)$.

PART III
THE GEODESIC BOUNDARY

In this part, we adapt to $\mathscr{W}_2(X)$ the classical construction of the cone topology on Hadamard space, see for instance [**Bal95**]. We introduce this

topology on $\overline{\mathscr{W}_2(X)} = \mathscr{W}_2(X) \cup \partial\mathscr{W}_2(X)$. We shall prove in Proposition 8.1 that the cone topology turns $\overline{\mathscr{W}_2(X)}$ into a first-countable Hausdorff space and that the topology induced on $\mathscr{W}_2(X)$ coincides with the topology derived from the Wassertein metric.

In the second part, we shall show in Theorem 9.2 that $\partial\mathscr{W}_2(X)$ is homeomorphic to

$$\mathscr{P}_1(c\partial X) = \left\{\zeta \in \mathscr{P}(c\partial X); \int s^2\,\zeta(d\xi, ds) = 1\right\}$$

endowed with the weak topology. In Corollary 9.3, we rewrite the above result in terms of Wassertein space over $c\partial X$.

Throughout this part, we will use the following notations. Let (Y, d) be a geodesic space and $y \in Y$. We set $\mathscr{R}_y(Y)$ (respectively $\mathscr{R}_{y,1}(Y)$) the set of geodesic rays in Y starting at y (respectively the set of unitary geodesic rays starting at y). These sets are closed subsets of $\mathscr{R}(Y)$ endowed with the topology of uniform convergence on compact subsets.

8. Cone topology on $\overline{\mathscr{W}_2(X)}$

The cone topology on $\overline{\mathscr{W}_2(X)} = \mathscr{W}_2(X) \cup \partial\mathscr{W}_2(X)$ is defined by using as a basis the open sets of $\mathscr{W}_2(X)$ together with

$$U(x, \xi, R, \varepsilon) = \left\{\theta \in \overline{\mathscr{W}_2(X)}; \theta \notin \overline{B}(\delta_x, R), \mathrm{W}((\mu_{\delta_x,\theta})_R, (\mu_{\delta_x,\xi})_R) < \varepsilon\right\}$$

where $x \in X$ is a fixed point, ξ runs over $\partial\mathscr{W}_2(X)$, R and ε run over $(0, +\infty)$ and $\mu_{\delta_x,\theta}$ is the unitary geodesic between δ_x and θ (existence and uniqueness follow from Lemma 8.3).

The main properties of the cone topology on $\overline{\mathscr{W}_2(X)}$ are gathered together in the following proposition.

Proposition 8.1. — *The cone topology on $\overline{\mathscr{W}_2(X)}$ is well-defined and is independent of the choice of the basepoint δ_x. Moreover, endowed with this topology, $\overline{\mathscr{W}_2(X)}$ is a first-countable Hausdorff space. By definition, the topology induced on $\mathscr{W}_2(X)$ coincides with the topology derived from the Wassertein metric.*

Remark 8.2. — We emphasize that the topology induced on $\partial\mathscr{W}_2(X)$ by the cone topology coincides with the quotient topology induced by the topology of uniform convergence on compact subsets on the set of unitary rays in $\mathscr{W}_2(X)$.

Moreover, since $\partial \mathscr{W}_2(X)$ endowed with the cone topology is first-countable, continuity and sequential continuity are equivalent in this topological space.

The scheme of proof is the same as in the nonpositively curved case. However, to get the result, we first need to generalize to our setting some properties related to nonpositive curvature.

Lemma 8.3. — *Given $x \in X$, the set of unitary rays in $\mathscr{W}_2(X)$ starting at δ_x is in one-to-one correspondence with the set $\mathscr{P}_1(c\partial X)$. Moreover, for any $\xi \in \partial \mathscr{W}_2(X)$, there exists a unique unitary ray starting at δ_x and belonging to ξ.*

Proof. — Recall that there exists a unique transport plan between a Dirac mass and any measure in $\mathscr{W}_2(X)$. Since there is a unique geodesic between two given points in X, the same property remains true for dynamical transportation plans. Using the previous remarks, we get that any $\mu \in \mathscr{P}(\mathscr{R}_x(X))$ such that $\int s^2(\gamma)\mu(d\gamma) < +\infty$ induces a ray starting at δ_x. Moreover, since displacement interpolation always exists (see Proposition 4.2), the set $\mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X))$ is in one-to-one correspondence with the unitary dynamical transportation plans starting at δ_x, namely the measures $\mu \in \mathscr{P}(\mathscr{R}_x(X))$ such that $\int s^2(\gamma)\mu(d\gamma) = 1$. Now, since X is a Hadamard space, we recall that $\mathscr{R}(X)$ is homeomorphic to $X \times c\partial X$ (where the left coordinate is the initial location of the ray). Therefore, for any $x \in X$, the previous map induces a homeomorphism

$$\phi_x : \mathscr{R}_x(X) \longrightarrow c\partial X. \tag{7}$$

This gives us a one-to-one correspondence between the set of unitary dynamical transportation plans starting at δ_x and the set $\mathscr{P}_1(c\partial X)$.

Given $\xi \in \partial \mathscr{W}_2(X)$, consider a unit ray (μ_t) in $\mathscr{W}_2(X)$ belonging to ξ and $\mu_\infty \in \mathscr{P}_1(c\partial X)$ its asymptotic measure. We claim that ${\phi_x^{-1}}_\#\mu_\infty$ is the unique ray starting at δ_x and belonging to ξ. Indeed, ${\phi_x^{-1}}_\#\mu_\infty$ and (μ_t) have the same asymptotic measure, thus they are asymptotic thanks to the asymptotic formula (Theorem 5.2). The asymptotic formula also implies that two asymptotic rays (starting at δ_x) have the same asymptotic measure, thus they are equal by the first part of the lemma. $\square$

Lemma 8.4. — *Let $(\mu_t), (\sigma_t)$ be two unitary geodesics (possibly rays) in $\mathscr{W}_2(X)$ starting at δ_x. Then, for any nonnegative s,t, the comparison angle $\widetilde{\angle}_{\mu_s\delta_x\sigma_t}$ at δ_x of the triangle $\Delta(\delta_x, \mu_s, \sigma_t)$ is a nondecreasing function*

of s and t. Consequently, the map $t \longrightarrow \mathrm{W}(\mu_t, \sigma_t)/t$ is a nondecreasing function as well.

Proof. — We set $d_m, d_s \leq +\infty$ the length of (μ_t) and (σ_t) respectively and μ, σ the corresponding optimal dynamical plans. Thanks to the lifting lemma, we set $\Theta \in \Gamma(\mu, \sigma)$ a dynamical plan such that, for given $s \leq d_m$ and $t \leq d_s$, $(e_s, e_t)_\# \Theta$ is an optimal plan. By definition of the Wassertein distance, we get, for any $s' \leq s$ and $t' \leq t$, the following estimate

$$\mathrm{W}^2(\mu_{s'}, \sigma_{t'}) \leq \int d^2(\gamma(s'), \gamma'(t'))\, \Theta(d\gamma, \gamma').$$

Now, the fact that X is nonpositively curved yields

$$\begin{aligned} d^2(\gamma(s'), \gamma'(t')) \ &\leq \ \frac{s'^2}{s^2} d^2(x, \gamma(s)) + \frac{t'^2}{t^2} d^2(x, \gamma'(t)) \\ &\quad - 2\frac{s't'}{st} d(x, \gamma(s)) d(x, \gamma'(t)) \cos \widetilde{\angle}_{\gamma(s) x \gamma'(t)}. \end{aligned}$$

where $\widetilde{\angle}_{\gamma(s) x \gamma'(t)}$ is the comparison angle at x (here, we use the fact that the initial measure is a Dirac mass). By integrating this inequality against Θ, we get

$$\begin{aligned} &\mathrm{W}^2(\mu_{s'}, \sigma_{t'}) \leq \\ &\qquad s'^2 + t'^2 - 2\frac{s't'}{st} \int d(x, \gamma(s)) d(x, \gamma'(t)) \cos \widetilde{\angle}_{\gamma(s) x \gamma'(t)}\, \Theta(d\gamma, d\gamma'). \end{aligned}$$

We conclude by noticing that the inequality above is an equality when $s = s'$ and $t = t'$, so we get

$$\mathrm{W}^2(\mu_{s'}, \sigma_{t'}) \leq s'^2 + t'^2 - 2s't' \cos \widetilde{\angle}_{\mu_s \delta_x \sigma_t}$$

which is equivalent to the property $\widetilde{\angle}_{\mu_{s'} \delta_x \sigma_{t'}} \leq \widetilde{\angle}_{\mu_s \delta_x \sigma_t}$. The remaining statement follows readily. □

Lemma 8.5. — *Let (μ_t) be a unitary geodesic, possibly a ray, starting at δ_y and $\delta_x \neq \delta_y$. For any $\theta \in \mathscr{W}_2(X)$ such that $\theta \neq \delta_x, \delta_y$, the comparison angle at θ satisfies*

$$\cos \widetilde{\angle}_{\delta_x \theta \delta_y} = \frac{1}{\mathrm{W}(\delta_x, \theta)\, \mathrm{W}(\delta_y, \theta)} \int d(x, z) d(y, z) \cos \widetilde{\angle}_{xzy}\, \theta(dz).$$

Moreover, given two nonnegative numbers $s \neq t$, the following inequality holds for $0 < t < T$

$$\widetilde{\angle}_{\mu_0 \mu_t \delta_x} + \widetilde{\angle}_{\mu_T \mu_t \delta_x} \geq \pi.$$

Proof. — For any $z \in X$, the following equality holds

$$d^2(x,y) = d^2(x,z) + d^2(y,z) - 2d(x,z)d(y,z)\cos \widetilde{\angle}_{xzy}.$$

By integrating this inequality against θ, we get the first statement by definition of the comparison angle. Let μ be the unique optimal dynamical coupling that induces (μ_t). The first step of the proof is to get the equality below:

$$\cos \widetilde{\angle}_{\mu_T \mu_t \delta_x} = \frac{1}{\mathrm{W}(\mu_T,\mu_t)\,\mathrm{W}(\mu_t,\delta_x)} \int d(\gamma(t),\gamma(T))d(\gamma(t),x)\cos \widetilde{\angle}_{x\gamma(t)\gamma(T)}\,\mu(d\gamma) \tag{8}$$

For $\gamma \in \operatorname{supp} \mu$, the following equality holds

$$d^2(\gamma(T),x) = d^2(\gamma(t),x) + d^2(\gamma(t),\gamma(T)) - 2d(\gamma(t),x)d(\gamma(t),\gamma(T))\cos \widetilde{\angle}_{x\gamma(t)\gamma(T)}.$$

By integrating this equality against μ, we get (8). Now, using that X is nonpositively curved, we have $\widetilde{\angle}_{x\gamma(t)\gamma(T)} + \widetilde{\angle}_{x\gamma(t)y} \geq \pi$. This gives

$$\begin{aligned}
&\cos \widetilde{\angle}_{\mu_T \mu_t \delta_x}\\
&\quad\leq \tfrac{-1}{\mathrm{W}(\mu_T,\mu_t)\,\mathrm{W}(\mu_t,\delta_x)} \int d(\gamma(t),\gamma(T))d(\gamma(t),x)\cos \widetilde{\angle}_{x\gamma(t)y}\,\mu(d\gamma)\\
&\quad\leq \tfrac{-1}{(T-t)\,\mathrm{W}(\mu_t,\delta_x)} \int \tfrac{T-t}{t} d(\gamma(t),y)d(\gamma(t),x)\cos \widetilde{\angle}_{x\gamma(t)y}\,\mu(d\gamma)\\
&\quad\leq \tfrac{-1}{\mathrm{W}(\mu_t,\delta_y)\,\mathrm{W}(\mu_t,\delta_x)} \int d(\gamma(t),y)d(\gamma(t),x)\cos \widetilde{\angle}_{x\gamma(t)y}\,\mu(d\gamma)\\
&\quad\leq -\cos \widetilde{\angle}_{\mu_0 \mu_t \delta_x}
\end{aligned}$$

where the last inequality follows from the first statement and the result is proved. □

As a consequence, we get the following result.

Proposition 8.6. — *Given $\varepsilon > 0$, $a > 0$, and $R > 0$, there exists a constant $T = T(\varepsilon, a, R) > 0$ such that the followings holds: for any $x, y \in X$ such that $d(x,y) = a$ and a unitary geodesic (possibly a ray) (μ_t) of length greater than T and starting at δ_y, if (σ_t^s) is the unitary geodesic from δ_x to μ_s then*

$$\mathrm{W}(\sigma_R^s, \sigma_R^{s'}) < \varepsilon$$

for any $s' > s > T$.

In particular, if (μ_t) is a ray and s goes to infinity, $(\sigma^s)_{s\geq 0}$ converges uniformly on compact subsets to the unitary ray $\mu_{\delta_x,\xi}$ where ξ is the asymptote class of (μ_t).

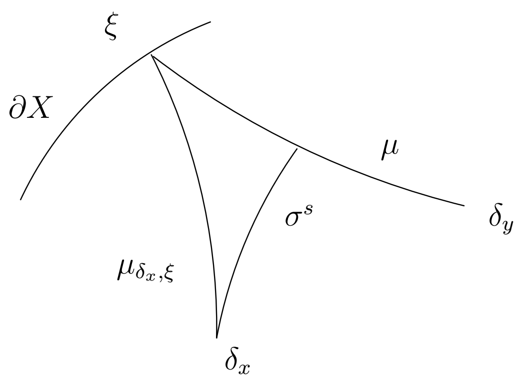

FIGURE 8. Uniform convergence of σ^s on compact subsets.

Proof. — Under these assumptions, the comparison angle $\widetilde{\angle}_{\delta_y\mu_s\delta_x}$ is arbitrary small provided s is sufficiently large. Consequently, thanks to Lemma 8.5, $\widetilde{\angle}_{\mu_{s'}\mu_s\delta_x}$ is close to π; therefore the comparison angle $\widetilde{\angle}_{\mu_s\delta_x\mu_{s'}}$ is small. This gives the first part of the result since

$$\widetilde{\angle}_{\sigma^s_R\delta_x\sigma^{s'}_R} \leq \widetilde{\angle}_{\mu_s\delta_x\mu_{s'}}$$

thanks to Lemma 8.4. Using Lemma 8.4 again, it only remains to prove the pointwise convergence of $(\sigma^s)_{s\geq 0}$ to $\mu_{\delta_x,\xi}$. Thanks to the asymptotic formula, there exists $C > 0$ such that

$$\mathrm{W}((\mu_{\delta_x,\xi})_t, \mu_t) \leq C \tag{9}$$

for any nonnegative number t. Finally, we conclude by using s' sufficiently large and the bound

$$\mathrm{W}((\mu_{\delta_x,\xi})_R, \sigma^s_R) \leq \mathrm{W}((\mu_{\delta_x,\xi})_R, \sigma^{s'}_R) + \mathrm{W}(\sigma^s_R, \sigma^{s'}_R)$$

where the same reasoning as above and (9) show that the first term on the right-hand side is small provided s' is large. □

Now, we can prove that the topology above is well-defined and does not depend on the choice of the base point δ_x. This the content of the lemma below.

Lemma 8.7. — *Given two positive numbers R, ε and $y \in X$, $\xi \in U(x,\eta,R,\varepsilon) \cap \partial \mathscr{W}_2(X)$, there exists $S, \varepsilon' > 0$ such that*

$$U(y,\xi,S,\varepsilon') \subset U(x,\eta,R,\varepsilon).$$

Proof. — We set $\alpha = \varepsilon - \mathrm{W}((\mu_{\delta_x,\eta})_R, (\mu_{\delta_x,\xi})_R) > 0$. Let $\theta \in U(y,\xi,S,\varepsilon')$ and Θ (respectively Ξ) be the unitary geodesic $\mu_{\delta_y,\theta}$ (respectively the unitary ray $\mu_{\delta_y,\xi}$). We have

$$\begin{aligned}\mathrm{W}((\mu_{\delta_x,\theta})_R, (\mu_{\delta_x,\eta})_R) \leq & \\ & \mathrm{W}((\mu_{\delta_x,\theta})_R, (\mu_{\delta_x,\Theta_S})_R) + \mathrm{W}((\mu_{\delta_x,\Theta_S})_R, (\mu_{\delta_x,\Xi_S})_R) \\ & + \mathrm{W}((\mu_{\delta_x,\Xi_S})_R, (\mu_{\delta_x,\xi})_R) + \mathrm{W}((\mu_{\delta_x,\xi})_R, (\mu_{\delta_x,\eta})_R)\end{aligned}$$

The first and the third term on the right-hand side are smaller than $\alpha/3$ for large S thanks to Proposition 8.6 while the second term is smaller than $\alpha/3$ for large S and small ε' thanks to lemma 8.4. □

9. The boundary of $\mathscr{W}_2(X)$ viewed as a set of measures

To state the main result of this part, we first need to introduce a definition.

Definition 9.1. — We set

$$\begin{array}{rccc} Am: & \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X)) & \longrightarrow & \mathscr{P}_1(c\partial X) \\ & (\mu_t) & \longmapsto & \mu_\infty \end{array}$$

the map that send a unitary ray starting at δ_x to its asymptotic measure.

The main result of this part is the following theorem.

Theorem 9.2. — *The map $Am : \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X)) \longrightarrow \mathscr{P}_1(c\partial X)$ induces a homeomorphism from $\partial \mathscr{W}_2(X)$ onto $\mathscr{P}_1(c\partial X)$.*

Note that a straightforward consequence of the result above is

Corollary 9.3. — *Let d be a metric on ∂X that induces the cone topology on ∂X and d_C the cone metric induced by d on $c\partial X$ (see (2) for a definition). Let us denote by $\mathscr{W}_2(c\partial X)$ the quadratic Wassertein space over the Polish space $(c\partial X, d_C)$. Then, $\partial \mathscr{W}_2(X)$ is homeomorphic to the subset of probability measures with unitary speed in $\mathscr{W}_2(c\partial X)$.*

Remark 9.4. — In particular we get the more symmetric result that $c\partial\mathscr{W}_2(X)$ is homeomorphic to $\mathscr{W}_2(c\partial X)$.

The rest of this part is devoted to the proof of the theorem above. Recall that we have proved in Lemma 8.3 that both Am and the map $\widetilde{Am} : \partial\mathscr{W}_2(X) \to \mathscr{P}_1(c\partial X)$ it induces are bijective.

The proof of Theorem 9.2 is in two steps. First, we prove that the map Am is a homeomorphism. Then, we use this fact to prove that $\widetilde{Am}$ is a homeomorphism as well.

We start the proof with a definition.

Definition 9.5. — Let $x \in X$. We denote by

$$ODT_x \quad = \quad \Big\{\mu \in \mathscr{P}(\mathscr{R}_x(X)); \int s^2(\gamma)\mu(d\gamma) = 1\Big\}$$

the set of unitary dynamical transport plans endowed with the weak topology. We also set

$$\begin{array}{rccl} (e_t)_{\geq 0\#} : & ODT_x & \longrightarrow & \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X)) \\ & \mu & \longmapsto & (\mu_t) \end{array}$$

and

$$\begin{array}{rccl} e_{\infty\#} : & ODT_x & \longrightarrow & \mathscr{P}_1(c\partial X) \\ & \mu & \longmapsto & \phi_{x\#}\mu \end{array}$$

where ϕ_x is defined in (7).

Thanks to Lemma 8.3, we have the following commutative diagram where all the maps are one-to-one.

$$\begin{tikzcd} ODT_x \arrow[r, "e_{\infty\#}"] \arrow[d, "(e_t)_{t\geq 0\#}"'] & \mathscr{P}_1(c\partial X) \\ \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X)) \arrow[ur, "Am"] & \end{tikzcd}$$

We first prove that

Lemma 9.6. — *The map $e_{\infty\#}$ is a homeomorphism onto $\mathscr{P}_1(c\partial X)$.*

Proof. — The map $\phi_x : \mathscr{R}_x(X) \longrightarrow c\partial X$ is a homeomorphism. Therefore it induces a homeomorphism between ODT_x and $\mathscr{P}_1(c\partial X)$ when endowed with the weak topology. $\square$

Lemma 9.7. — *The map $(e_t)_{\geq 0\#}$ is a continuous map.*

Proof. — Since the spaces we consider are metrizable, we just have to prove the sequential continuity. Consequently, we are given a sequence $(\mu_n)_{n\in\mathbb{N}}$ such that $\mu_n \rightharpoonup \mu$ in ODT_x. Now, since $e_t : \mathscr{R}_x(X) \longrightarrow X$ is a continuous map, we get that $e_{t\#}\mu_n \rightharpoonup e_{t\#}\mu$ in $\mathscr{P}(X)$. By definition of ODT_x, we have

$$\int s^2(\gamma)\,\mu_n(d\gamma) = \int s^2(\gamma)\,\mu(d\gamma) = 1.$$

Since $\int s^2(\gamma)\,\mu_n(d\gamma) = \int d^2(x,\gamma(1))\,\mu_n(d\gamma)$, the equality above implies the convergence of the second moment. Namely, we have

$$\int d^2(x,\gamma(t))\,\mu_n(d\gamma) = t^2 \int d^2(x,\gamma(1))\,\mu_n(d\gamma) = \int d^2(x,\gamma(t))\,\mu(d\gamma).$$

This implies the convergence of $e_{t\#}\mu_n$ to $e_{t\#}\mu$ with respect to the Wassertein distance (see for instance [**Vil09**]). Thus, we have proved the pointwise convergence of rays. Now, since $t \longrightarrow \mathrm{W}(e_{t\#}\mu_n, e_{t\#}\mu)$ is nondecreasing as proved in Lemma 8.4, we get the result. □

We end the first part of the proof with the following lemma.

Lemma 9.8. — *The map $(e_t)_{\geq 0\#}$ is a homeomorphism.*

Proof. — Since the topology of both $\mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X))$ and ODT_x is induced by a metric, it is sufficient to prove that $(e_t)_{\geq 0\#}$ is a proper map. Moreover, we just have to prove sequential compactness. We set K a compact subset of $\mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X))$. Let $(\mu_n)_{n\in\mathbb{N}} \in (e_t)_{\geq 0\#}^{-1}(K)$. We first notice that $(\mu_n)_{n\in\mathbb{N}}$ is tight. Indeed, by assumption on K, the sequence $(e_{1\#}\mu_n)_{n\in\mathbb{N}}$ is tight in $\mathscr{P}(X)$. Therefore, by arguing as in the end of the proof of Proposition 4.2, we obtain the claim. Consequently, since $\mathscr{R}_x(X)$ is a Polish space, we can apply Prokhorov's theorem to get a converging subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ to $\widetilde{\mu}$. It remains to prove that $\widetilde{\mu} \in ODT_x$, namely that $\int s^2(\gamma)\,\widetilde{\mu}(d\gamma) = 1$. Since K is compact, we can also assume without loss of generality that $(e_{t\#}\mu_{n_k}) \longrightarrow (\bar{\mu}_t)$ in $\mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X))$. Moreover, since $e_{t\#}\mu_{n_k} \rightharpoonup e_{t\#}\widetilde{\mu}$ for any t, we get $(e_t)_{\geq 0\#}(\widetilde{\mu}) = (\bar{\mu}_t)$. Therefore, $e_{1\#}\mu_{n_k} \longrightarrow e_{1\#}\widetilde{\mu}$ in $\mathscr{W}_2(X)$. This implies the convergence of the second moment $\int d^2(x,\gamma(1))\,\mu_{n_k}(d\gamma) = \int s^2(\gamma)\,\mu_{n_k}(d\gamma) = 1$ (see for instance [**Vil09**]) and the result is proved. □

We are now in position to prove Theorem 9.2. We set $p_{\partial\mathscr{W}_2}$ the canonical projection on $\partial\mathscr{W}_2(X)$. We have the following commutative diagram.

$$\begin{tikzcd}
\mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X)) \arrow[r, two heads, "Am"] \arrow[d, two heads, "p_{\partial\mathscr{W}_2}"'] & \mathscr{P}_1(c\partial X) \\
\partial\mathscr{W}_2(X) \arrow[ur, two heads, "\widetilde{Am}"] &
\end{tikzcd}$$

We have seen at the beginning of the proof that all the maps above are one-to-one. To conclude, it remains to prove that $p^{-1}_{\partial\mathscr{W}_2}$ is a continuous map. Since $\partial\mathscr{W}_2(X)$ is first-countable (see Remark 8.2), it is sufficient to prove sequential continuity. To this aim, let $\xi_n \longrightarrow \xi$ in $\partial\mathscr{W}_2(X)$ and $(e_{t\#}\mu_n), (\mu_t) \in \mathscr{R}_{\delta_x,1}(\mathscr{W}_2(X))$ such that $p_{\partial\mathscr{W}_2}((e_{t\#}\mu_n)) = \xi_n$ and $p_{\partial\mathscr{W}_2}((\mu_t)) = \xi$. Recall that under these assumptions, the map $t \longrightarrow \mathrm{W}(e_{t\#}\mu_n, \mu_t)$ is nondecreasing (see Lemma 8.4), thus we just have to show the pointwise convergence of $(e_{t\#}\mu_n)$. This pointwise convergence follows readily from the definition of the cone topology on $\overline{\mathscr{W}_2(X)}$.

PART IV
ISOMETRIC RIGIDITY

10. Statement of the result and reduction to a Radon transform

In this part we study the isometry group of $\mathscr{W}_2(X)$. We have a natural morphism

$$\begin{aligned} \# : \operatorname{Isom} X &\to \operatorname{Isom} \mathscr{W}_2(X) \\ \varphi &\mapsto \varphi_\#. \end{aligned}$$

An isometry Φ of $\mathscr{W}_2(X)$ is said to be *trivial* if it is in the image of #, to *preserve shape* if for all $\mu \in \mathscr{W}_2(X)$ there is an isometry φ (depending on μ) of X such that $\Phi(\mu) = \varphi_\#\mu$, and to be *exotic* otherwise. When all isometries of $\mathscr{W}_2(X)$ are trivial, that is when # is an isomorphism, we say that $\mathscr{W}_2(X)$ is *isometrically rigid*.

The main result of this part is the following.

***Theorem 10.1* (isometric rigidity).** — *Let X be a geodesically complete, negatively curved Hadamard space. If*

(1) *X is a tree, or*

(2) *X is a Riemannian manifold, or more generally*

(3) *there is a dense subset of X all of whose points admit a neighborhood isometric to an open subset of a Riemannian manifold of dimension at least* 2*,*

then $\mathscr{W}_2(X)$ is isometrically rigid.

By "Negatively curved" we mean that X is not a line and the CAT(0) inequality is strict except for triangles all of whose vertices are aligned. Note that when X is Euclidean, $\mathscr{W}_2(X)$ is not isometrically rigid as proved in [**Klo08**]. The $I_{p,q}$ building, as well as many polyhedral complexes whose faces are endowed with hyperbolic metrics, provide examples of isometric rigidity covered by the third case of the Theorem.

The strategy of proof is as follows: we show first that an isometry must map Dirac masses to Dirac masses. The method is similar to that used in the Euclidean case, and is valid for geodesically complete locally compact spaces without curvature assumption. It already yields Corollary 10.5, answering positively in a wide setting the following question: if $\mathscr{W}_2(Y)$ and $\mathscr{W}_2(Z)$ are isometric, can one conclude that Y and Z are?

The second step is to prove that an isometry that fixes all Dirac masses must map a measure supported on a geodesic to a measure supported on the same geodesic. Here we use that X has negative curvature: this property is known to be false in the Euclidean case.

Then as in [**Klo08**] we are reduced to prove the injectivity of a specific Radon transform. The point is that this injectivity is known for $\mathbb{R}\mathrm{H}^n$ and is easy to prove for trees, but seems not to be known for other spaces. We give a partial but sufficient injectivity argument for manifolds. This argument can be generalized to spaces with a dense set of higher-dimensional manifold-type points.

10.1. Characterization of Dirac masses. — The characterization of Dirac masses follows from two lemmas, and can be carried out without any assumption on the curvature. Note that when we ask a space to be geodesically complete, we of course imply that it is geodesic.

Lemma 10.2. — *Let Y be a locally compact, geodesically complete Polish space. Any geodesic segment in $\mathscr{W}_2(Y)$ issued from a Dirac mass can be extended to a complete geodesic ray.*

Proof. — Let us recall the measurable selection theorem (see for example [**Del75**], Corollary of Theorem 17): any surjective measurable map between Polish spaces admits a measurable right inverse provided its fibers are compact. Consider the restriction map

$$p^T : \mathscr{R}(Y) \to \mathscr{G}^{0,T}(Y) \ ;$$

the fiber of a geodesic segment γ is closed in the set $\mathscr{R}_{\gamma_0,s}(Y)$ of geodesic rays of speed $s = s(\gamma)$ starting at γ_0. This set is compact by Arzela-Ascoli theorem (equicontinuity follows from the speed being fixed, while the pointwise relative compactness is a consequence of Y being locally compact and geodesic, hence proper).

Moreover the geodesic completeness implies the surjectivity of p^T. There is therefore a measurable right inverse q of p^T: it maps a geodesic segment γ to a complete ray whose restriction to $[0,T]$ is γ.

Let $(\mu_t)_{t\in[0,T]}$ be a geodesic segment of $\mathscr{W}_2(Y)$, of speed s, with $\mu_0 = \delta_x$ and let $\mu = q_\#(\mu_T)$: it is a probability measure on $\mathscr{R}(Y)$ whose restriction $p^T_\#(\mu)$ is the displacement interpolation of (μ_t). For all $t > 0$ we therefore denote by μ_t the measure $(e_t)_\#\mu$ on X.

It is easy to see that it defines a geodesic ray: since there is only one transport plan from a Dirac mass to any fixed measure, $\mathrm{W}^2(\mu_0,\mu_t) = \int s(\gamma)^2 t^2\,\mu(d\gamma) = s^2t^2$. Moreover the transport plan from μ_t to $\mu_{t'}$ deduced from μ gives $\mathrm{W}(\mu_t,\mu_{t'}) \leqslant s|t-t'|$. But the triangular inequality applied to μ_0, μ_t and $\mu_{t'}$ implies $\mathrm{W}(\mu_t,\mu_{t'}) \geqslant s|t-t'|$ and we are done. □

Lemma 10.3. — *If Y is a geodesic Polish space, given $\mu_0 \in \mathscr{W}_2(Y)$ not a Dirac mass and $y \in \operatorname{supp}\mu_0$, the geodesic segment from μ_0 to $\mu_1 := \delta_y$ cannot be geodesically extended for any time $t > 1$.*

Note that at least in the branching case the assumption $y \in \operatorname{supp}\mu_0$ is needed.

Proof. — Assume that there is a geodesic segment $(\mu_t)_{t\in[0,1+\varepsilon]}$ and let μ be an optimal dynamical transport plan from μ_0 to $\mu_{1+\varepsilon}$. Since $y \in \operatorname{supp}\mu_0$, $\operatorname{supp}\mu$ contains a geodesic segment of X that is at y at times 0 and 1, and must therefore be constant.

Since μ_0 is not a Dirac mass, there is a $y' \neq y$ in $\operatorname{supp}\mu_0$. Let γ be a geodesic segment in $\operatorname{supp}\mu$ such that $\gamma_0 = y'$. Then $\gamma_1 = y$ lies between

y' and $y'' := \gamma_{1+\varepsilon}$ on γ, and the transport plan $\Pi = \mu^{0,1+\varepsilon}$ contains in its support the couples (y'', y') and (y, y).

But then cyclical monotonicity shows that Π is not optimal: it costs less to move y'' to y and y to y' by convexity of the cost (see figure 9). The dynamical transport μ cannot be optimal either, a contradiction. $\square$

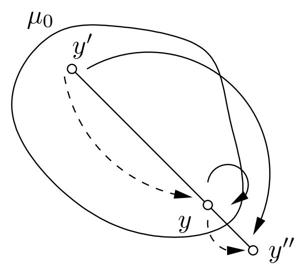

FIGURE 9. The transport shown with continuous arrows is less effective than the transport given by the dashed arrows.

We can now easily draw the consequences of these lemmas.

Proposition 10.4. — *Let Y, Z be geodesic Polish spaces and assume that Y is geodesically complete and locally compact. Any isometry from $\mathscr{W}_2(Y)$ to $\mathscr{W}_2(Z)$ must map all Dirac masses to Dirac masses.*

Except in the next corollary, we shall use this result with $Y = Z = X$. Note that we will need to ask that X be geodesically complete in addition to the Hadamard hypothesis.

Proof. — Denote by φ an isometry $\mathscr{W}_2(Y) \to \mathscr{W}_2(Z)$ and consider any $x \in Y$. If $\varphi(\delta_x)$ were not a Dirac mass, there would exist a geodesic segment $(\mu_t)_{t\in[0,1]}$ from $\varphi(\delta_x)$ to a Dirac mass (at a point $y \in \operatorname{supp}\varphi(\delta_x)$) that cannot be extended for times $t > 1$.

But $\varphi^{-1}(\mu_t)$ gives a geodesic segment issued from δ_x, that can therefore be extended. This is a contradiction since φ is an isometry. $\square$

Corollary 10.5. — *Let Y, Z be geodesic Polish spaces and assume that Y is geodesically complete and locally compact. Then $\mathscr{W}_2(Y)$ is isometric to $\mathscr{W}_2(Z)$ if and only if Y is isometric to Z.*

Proof. — Let φ be an isometry $\mathscr{W}_2(Y) \to \mathscr{W}_2(Z)$. Then φ maps Dirac masses to Dirac masses, and since the set of Dirac masses of a space is

canonically isometric to the space, φ induces an isometry $Y \to Z$. the converse implication is obvious. □

Note that we do not know whether this result holds for general metric spaces. Also, we do not know if there is a space Y and an isometry of $\mathscr{W}_2(Y)$ that maps some Dirac mass to a measure that is not a Dirac mass.

10.2. Measures supported on a geodesic. — The characterization of measures supported on a geodesic relies on the following argument: when dilated from a point of the geodesic, such a measure has Euclidean expansion.

Lemma 10.6. — *Assume that X is negatively curved, and let γ be a maximal geodesic of X, μ be in $\mathscr{W}_2(X)$. Given a point $x \in \gamma$, denote by $(x^t \cdot \mu)_{t\in[0,1]}$ the geodesic segment from δ_x to μ. If y is any point of X, $(x+y)/2$ denotes the midpoint of x and y.*

The measure μ is supported on γ if and only if for all $x, g \in \gamma$

$$\mathrm{W}(x^{\frac{1}{2}} \cdot \mu, x^{\frac{1}{2}} \cdot \delta_g) = \frac{1}{2}\mathrm{W}(\mu, \delta_g).$$

Proof. — The "only if" part is obvious since the transport problem on a convex subset of X only involves the induced metric on this subset, which here is isometric to an interval.

To prove the "if" part, first notice that the CAT(0) inequality in a triangle (x, y, g) yields the Thales inequality $d((x + y)/2, (x + g)/2) \leqslant \frac{1}{2}d(y, g)$, so that by direct integration

$$\mathrm{W}(x^{\frac{1}{2}} \cdot \mu, \delta_{(x+g)/2}) \leqslant \frac{1}{2}\mathrm{W}(\mu, \delta_g). \tag{10}$$

Note also that $x^{\frac{1}{2}} \cdot \delta_g = \delta_{(x+g)/2}$.

Assume now that μ is not supported on γ and let $y \in \operatorname{supp}\mu \setminus \gamma$, and $x, g \in \gamma$ such that x, y, g are not aligned. Since γ is maximal, this is possible, for example by taking $d(x, g) > d(x, y)$ in the branching case (any choice of $x \neq g$ would do in the non-branching case). Since X is negatively curved, we get that $d((x, y)/2, (x + g)/2) < \frac{1}{2}d(y, g)$ so that (10) is strict (see Figure 10). □

Corollary 10.7. — *If X is a negatively curved Hadamard space and γ is a maximal geodesic of X, any isometry of $\mathscr{W}_2(X)$ that fixes Dirac masses must preserve the subset $\mathscr{W}_2(\gamma)$ of measures supported on γ, and therefore induces an isometry on this set.*

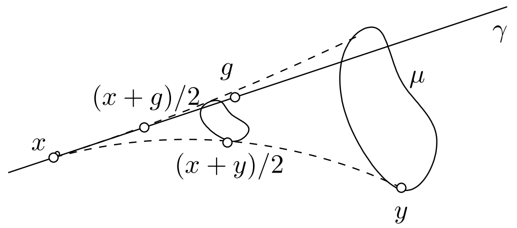

Figure 10. In negative curvature, midpoinds are closer than they would be in Euclidean space.

Proof. — Let φ be an isometry of $\mathscr{W}_2(X)$ that fixes Dirac masses. Let γ be a maximal geodesic of X and $\mu \in \mathscr{W}_2(X)$ be supported on γ. If $\varphi(\mu)$ where not supported in γ, there would exist $x, g \in \gamma$ such that $\mathrm{W}(\varphi(\mu)_{1/2}, \delta_{(x+g)/2}) < \frac{1}{2}\mathrm{W}(\varphi(\mu), \delta_g)$. But φ is an isometry and $\varphi^{-1}(\delta_x) = \delta_x$, $\varphi^{-1}(\delta_g) = \delta_g$, $\varphi^{-1}(\delta_{(x+g)/2}) = \delta_{(x+g)/2}$ so that also $\varphi(\mu)_{1/2} = \varphi(\mu_{1/2})$. As a consequence, (10) would be a strict inequality too, a contradiction. □

10.3. Isometry induced on a geodesic. — We want to deduce from the previous section that an isometry that fixes all Dirac masses must also fix every geodesically-supported measure. To this end, we have to show that the isometry induced on the measures supported in a given geodesic is trivial, in other terms to rule out the other possibilities exhibited in [**Klo08**] at which it is recommanded to take a look before reading the proof below.

Proposition 10.8. — *Assume that X is a geodesically complete, negatively curved Hadamard space and let φ be an isometry of $\mathscr{W}_2(X)$ that fixes all Dirac masses. For all complete geodesics γ of X, the isometry induced by φ on $\mathscr{W}_2(\gamma)$ is the identity.*

We only adress the geodesically complete case for simplicity, but the same probably holds in more generality.

Proof. — Let x be a point not lying on γ. Such a points exists since X is negatively curved, hence not a line.

First assume that φ induces on $\mathscr{W}_2(\gamma)$ an exotic isometry. Let y, z be two points of γ and define $\mu_0 = \frac{1}{2}\delta_y + \frac{1}{2}\delta_z$. Then $\mu_n := \varphi^n(\mu_0)$ has the form $m_n\delta_{y_n} + (1-m_n)\delta_{z_n}$ where $m_n \to 0$, $y_n \to \infty$, $z_n \to (y+z)/2 =: g$.

Let now γ' be a complete geodesic that contains $y' = (x+y)/2$ and $z' = (x+z)/2$ (see Figure 11). Since the midpoint of μ_0 and δ_x is

supported on γ', so is the midpoint of μ_n and δ_x. This means that $(x+y_n)/2$ and $(x+z_n)/2$ lie on γ', and the former goes to infinity. This shows that φ also induces an exotic isometry on $\mathscr{W}_2(\gamma')$, thus that $(x+z_n)/2$ goes to the midpoint g' of y' and z'. But we already know that $(x+z_n)/2$ tends to the midpoint of x and g, which must therefore be g'.

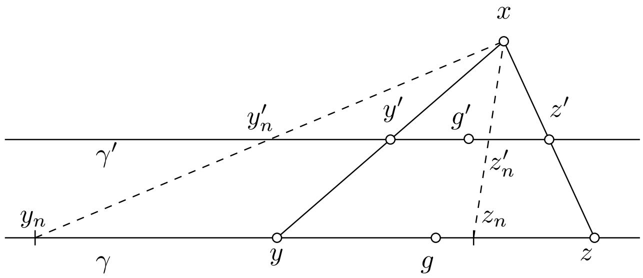

FIGURE 11. That an isometry acts exotically on a geodesic γ would imply that there is another geodesic γ' such that taking midpoints with x is an affine map. Then γ and γ' would be parallel, therefore ound a flat strip.

Since this holds for all choices of y and z, we see that the map $y \to (x+y)/2$ maps affinely γ to γ'. But the geodesic segment $[x\gamma_t]$ converges when $t \to \pm\infty$ to geodesic segments asymptotic to γ and $-\gamma$ respectively. It follows that γ' is parallel to γ, so that they must bound a flat strip (see [**Bal95**]). But this is forbidden by the negative curvature assumption.

A similar argument can be worked out in the case when φ induces an involution: given $y, z \in \gamma$ and their midpoint g, one can find measures μ_n supported on y and y_n, where $y_n \to g$ and μ_n has more mass on y_n than on y, such that $\varphi(\mu_n)$ is supported on z and a point z_n of γ, with more mass on z_n. It follows that the midpoints y', y'_n, z' with x are on a line, and we get the same contradiction as before.

The classification of isometries of $\mathscr{W}_2(\mathbb{R})$ shows that if φ fixes Dirac masses and is neither an exotic isometry nor an involution, then it is the identity. $\square$

Now we are able to link the isometric rigidity of $\mathscr{W}_2(X)$ to the injectivity of a Radon transform. The following definition relies on the following observation: since a geodesic is convex and X is Hadamard, given a point y and a geodesic γ there is a unique point $p_\gamma(y) \in \gamma$ closest to y, called the projection of y to γ.

Definition 10.9. — When X is geodesically complete, we define the *perpendicular Radon transform* $\mathscr{R}\mu$ of a measure $\mu \in \mathscr{W}_2(X)$ as the following map defined over complete geodesics γ of X:

$$\mathscr{R}\mu(\gamma) = (p_\gamma)_\#\mu.$$

In other words, this Radon transform recalls all the projections of a measure on geodesics.

The following is now a direct consequence of Proposition 10.8.

Corollary 10.10. — *Assume that X is geodesically complete and negatively curved. If there is a dense subset $A \subset \mathscr{W}_2(X)$ such that for all $\mu \in \mathscr{W}_2(X)$ and all $\nu \in A$, we have*

$$\mathscr{R}\mu = \mathscr{R}\nu \Rightarrow \mu = \nu$$

then $\mathscr{W}_2(X)$ is isometrically rigid.

Proof. — Let φ be an isometry of $\mathscr{W}_2(X)$. Up to composing with an element of the image of $\#$, we can assume that φ acts trivially on Dirac masses. Then it acts trivially on geodesically supported measures. For all $\nu \in A$, since $(p_\gamma)_\#\nu$ is the measure supported on γ closest to ν, one has $\mathscr{R}\varphi(\nu) = \mathscr{R}\nu$, hence $\varphi(\nu) = \nu$. We just proved that φ acts trivially on a dense set, so that it must be the identity. □

11. Injectivity of the Radon transform

The proof of Theorem 10.1 shall be complete as soon as we get the injectivity required in Corollary 10.10. Note that in the case of the real hyperbolic space $\mathbb{R}\mathrm{H}^n$, we could use the usual Radon transform on the set of compactly supported measures with smooth density to get it (see [**Hel99**]).

11.1. The case of manifolds and their siblings. — Let us first give an argument that does the job for all manifolds. We shall give a more general, but somewhat more involved argument afterward, and the next subsection shall deal with the case of trees.

Proposition 11.1. — *Assume that X is a Hadamard manifold. Let A be the set of finitely supported measures. For all $\mu \in \mathscr{W}_2(X)$ and all $\nu \in A$, if $\mathscr{R}\mu = \mathscr{R}\nu$ then $\mu = \nu$.*

Note that A is dense in $\mathscr{W}_2(X)$ so that this proposition ends the proof of Theorem 10.1, case 2.

Proof. — Write $\nu = \sum m_i \delta_{x_i}$ where $\sum m_i = 1$. Note that since X is a *negatively* curved manifold, it has dimension at least 2.

First, we prove under the assumption $\mathscr{R}\,\mu = \mathscr{R}\,\nu$ that μ must be supported on the x_i. Let x be any other point, and consider a geodesic γ such that $\gamma(0) = x$ and $\dot\gamma(0)$ is not orthogonal to any of the geodesics (xx_i). Then for all i, $p_\gamma(x_i) \neq x$ and there is an $\varepsilon > 0$ such that the neighborhood of size ε around x on γ does not contain any of these projections. It follows that $\mathscr{R}\,\nu(\gamma)$ is supported outside this neighborhood, and so does $\mathscr{R}\,\mu(\gamma)$. But the projection on γ is 1-Lipschitz, so that μ must be supported outside the ε neighborhood of x in X. In particular, $x \notin \operatorname{supp}\mu$.

Now, if γ is a geodesic containing x_i, then $\mathscr{R}\,\nu(\gamma)$ is finitely supported with a mass at least m_i at x_i. For a generic γ, the mass at x_i is exactly m_i. It follows immediately, since μ is supported on the x_i, that its mass at x_i is m_i. □

The above proof mainly uses the fact that given a point x and a finite number of other points x_i, there is a geodesic $\gamma \ni x$ such that $p_\gamma(x_i) \neq x$ for all i. It follows that the proof can be adjusted to get the case 3 of Theorem 10.1. To this end, let us define the *regular set* of X as the set of points having a neighborhood isometric to an open set in a Riemannian manifold of dimension at least 2.

Proposition 11.2. — *Assume that X is a Hadamard, geodesically complete space. Let A be the set of measures supported on finite set of points all located in the regular set of X. For all $\mu \in \mathscr{W}_2(X)$ and all $\nu \in A$, if $\mathscr{R}\,\mu = \mathscr{R}\,\nu$ then $\mu = \nu$.*

Proof. — Write $\nu = \sum m_i \delta_{x_i}$ where $\sum m_i = 1$. Consider one of the x_i, and let $m'_i := \mu(\{x_i\})$. We get $m'i \leq m_i$ from $\mathscr{R}\,\mu = \mathscr{R}\,\nu$.

Since by hypothesis, x_i has a neighborhood U_i isometric to an open set in a Riemannian manifold of dimension at least 2, as before there is a geodesic $\gamma \ni x_i$ such that $p_\gamma(x_j) \neq x_i$ for all $j \neq i$. It follows that for some $\varepsilon > 0$, the measure $\mathscr{R}\,\mu(\gamma) = \mathscr{R}\,\nu(\gamma)$ is concentrated outside $B(x_i, \varepsilon) \setminus x_i$. By taking ε small enough so that $B(x_i, \varepsilon)$ is contained in U_i, it follows also that μ is concentrated outside $B(x_i, \varepsilon) \setminus x_i$.

Let $\gamma^\perp$ denote the set of points x such that the geodesic segment $[xx_i]$ is orthogonal to γ at x_i. Then $\gamma^\perp \cap B(x_i, \varepsilon)$ is the exponential image of

the orthogonal of the tangent line of γ at x_i. By definition of the Radon transform, $\mu(\gamma^\perp \setminus x_i) = m_i - m_i'$ and for all $x \in \operatorname{supp}\mu \setminus \gamma^\perp$, we have $p_\gamma(x) \notin B(x_i, \varepsilon)$.

Choose a second geodesic $\gamma_2 \ni x_i$, close enough to γ to ensure that for all $x \in \operatorname{supp}\mu \setminus \gamma^\perp$, we still have $p_{\gamma_2}(x) \notin B(x_i, \varepsilon)$ (up to schrinking ε a bit if necessary). We can moreover assume that γ_2 enjoys the same properties we asked to γ, so that

$$\forall x \in \operatorname{supp}\mu \setminus \gamma^\perp \cap \gamma_2^\perp, \quad p_{\gamma_2}(x) \notin B(x_i, \varepsilon).$$

Let n be the dimension of U_i. We can construct inductively a family $\gamma_1, \dots, \gamma_n$ of geodesics chosen as above, and such that their velocity vectors at x_i span its tangent space. For all point x in

$$\operatorname{supp}\mu \setminus \cap_\alpha \gamma_\alpha^\perp,$$

we get $p_{\gamma_n}(x) \notin B(x_i, \varepsilon)$. But $\cap_\alpha \gamma_\alpha^\perp = \{x_i\}$, and considering $\mathscr{R}\mu(\gamma_n) = \mathscr{R}\nu(\gamma_n)$ we get $m_i' = m_i$. Since μ is a probability measure and $\sum m_i = 1$, we deduce $\mu = \nu$. □

Now if X fits in case 3 of Theorem 10.1, then the regular set of X is dense, thus A is dense in $\mathscr{W}_2(X)$. We can therefore conclude by Corollary 10.10.

11.2. The case of trees. — Let X be a locally finite tree that is not a line. Recall the notations already used in Section 6. We describe X by a couple (V, E) where V is the set of vertices; E is the set of edges, each endowed with one or two endpoints in V and a length.

For all $x \in V$, let $k(x)$ be the valency of x, that is the number of edges incident to x. We assume that no vertex has valency 2, since otherwise we could describe the same metric space by a simpler graph. Note also that edges with only one endpoint have infinite length since X is assumed to be complete.

In this setting, X is geodesically complete if and only if it has no leaf (vertex of valency 1). Assume this, and let γ be a complete geodesic. If x is a vertex lying on γ and $C_1, \dots, C_{k(x)}$ are the connected components of $X \setminus \{x\}$, let $\perp^x(\gamma)$ be the union of x and of the C_i not meeting γ (see Figure 12). It inherits a tree structure from X. In fact, $\perp^x(\gamma)$ depends only upon the two edges e, f of γ that are incident to x. We therefore let $\perp^x(ef) = \perp^x(\gamma)$.

The levels of p_γ are called the *perpendiculars* of γ, they also are the bissectors of its points. They are exactly:

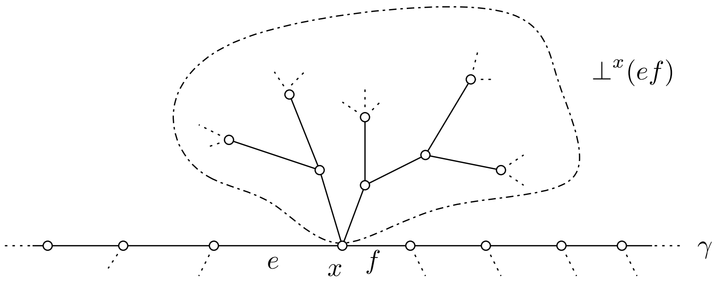

FIGURE 12. A perpendicular to a geodesic.

– the sets $\perp^x(\gamma)$ where x is a vertex of γ.
– the sets $\{x\}$ where x is a point interior to an edge of γ.

A measure $\mu \in \mathscr{W}_2(X)$ can be decomposed into a part supported outside vertices, which is obviously determined by the projections of μ on the various geodesics of X, and an atomic part supported on vertices. Therefore, we are reduced to study the perpendicular Radon transform reformulated as follows for functions defined on V instead of measures on X.

Definition 11.3 **(combinatorial Radon transform)**
A *flag* of X is defined as a triple (x, ef) where x is a vertex, $e \neq f$ are edges incident to x and ef denotes an unordered pair. Let us denote the set of flags by F; we write $x \in ef$ to say that $(x, ef) \in F$.

Given a summable function h defined on the vertices of X, we define its *combinatorial perpendicular Radon transform* as the map

$$\begin{array}{rccc} \mathscr{R}\,h: & F & \to & \mathbb{R} \\ & (x,ef) & \mapsto & \displaystyle\sum_{y\in\perp^x(ef)} h(y) \end{array}$$

where the sum is on vertices of $\perp^x(ef)$.

It seems that this Radon transform has not been studied before, contrary to the transforms defined using geodesics [**BCTCP91**], horocycles [**Car73, BFP89**] and circles [**CTGP03**].

Theorem 11.4 **(Inversion formula)**. — *Two maps $h, l : V \to \mathbb{R}$ such that $\sum h = \sum l$ and $\mathscr{R}\,h = \mathscr{R}\,l$ are equal. More precisely, we can recover*

h from $\mathscr{R}h$ by the following inversion formula:

$$h(x) = \frac{1}{k(x)-1}\sum_{ef\ni x}\mathscr{R}h(x,ef) - \frac{k(x)-2}{2}\sum_{y\in V}h(y)$$

where the first sum is over the set of pairs of edges incident to x.

Proof. — The formula relies on a simple double counting argument:

$$\begin{aligned}\sum_{ef\ni x}\mathscr{R}h(x,ef) &= \sum_{ef\ni x}\sum_{y\in\perp^x(ef)}h(y)\\ &= \sum_{y\in V}h(y)n_x(y)\end{aligned}$$

where $n_x(y)$ is the number of flags (x,ef) such that $y\in\perp^x(ef)$. If $y\neq x$, let e_y be the edge incident to x starting the geodesic segment from x to y. Then $y\in\perp^x(ef)$ if and only if $e,f\neq e_y$. Therefore, $n_x(y)=\binom{k(x)-1}{2}$. But $n_x(x)=\binom{k(x)}{2}$, so that

$$\sum_{ef\ni x}\mathscr{R}h(x,ef) = \binom{k(x)-1}{2}\sum_{y\in V}h(y) + (k(x)-1)h(x).$$

$\square$

References

[Bal95] W. Ballmann – *Lectures on spaces of nonpositive curvature*, DMV Seminar, vol. 25, Birkhäuser Verlag, Basel, 1995, With an appendix by Misha Brin.

[BCTCP91] C. A. Berenstein, E. Casadio Tarabusi, J. M. Cohen & M. A. Picardello – “Integral geometry on trees”, *Amer. J. Math.* **113** (1991), no. 3, p. 441–470.

[BFP89] W. Betori, J. Faraut & M. Pagliacci – “An inversion formula for the Radon transform on trees”, *Math. Z.* **201** (1989), no. 3, p. 327–337.

[BH99] M. R. Bridson & A. Haefliger – *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 319, Springer-Verlag, Berlin, 1999.

[Bil99] P. Billingsley – *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and

Statistics, John Wiley & Sons Inc., New York, 1999, A Wiley-Interscience Publication.

[Bou97] M. Bourdon – "Immeubles hyperboliques, dimension conforme et rigidité de Mostow", *Geom. Funct. Anal.* **7** (1997), no. 2, p. 245–268.

[BP99] M. Bourdon & H. Pajot – "Poincaré inequalities and quasiconformal structure on the boundary of some hyperbolic buildings", *Proc. Amer. Math. Soc.* **127** (1999), no. 8, p. 2315–2324.

[Bus55] H. Busemann – *The geometry of geodesics*, Academic Press Inc., New York, N. Y., 1955.

[Car73] P. Cartier – "Harmonic analysis on trees", Harmonic analysis on homogeneous spaces (Proc. Sympos. Pure Math., Vol. XXVI, Williams Coll., Williamstown, Mass., 1972), Amer. Math. Soc., Providence, R.I., 1973, p. 419–424.

[CTGP03] E. Casadio Tarabusi, S. G. Gindikin & M. A. Picardello – "The circle transform on trees", *Differential Geom. Appl.* **19** (2003), no. 3, p. 295–305.

[Del75] C. Dellacherie – "Ensembles analytiques. Théorèmes de séparation et applications", Séminaire de Probabilités, IX (Seconde Partie, Univ. Strasbourg, Strasbourg, années universitaires 1973/1974 et 1974/1975), Springer, Berlin, 1975, p. 336–372. Lecture Notes in Math., Vol. 465.

[EO73] P. Eberlein & B. O'Neill – "Visibility manifolds", *Pacific J. Math.* **46** (1973), p. 45–109.

[Hel99] S. Helgason – *The Radon transform*, second ed., Progress in Mathematics, vol. 5, Birkhäuser Boston Inc., Boston, MA, 1999.

[Klo08] B. Kloeckner – "A geometric study of wasserstein spaces: euclidean spaces", 2008, to appear in *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*.

[Lot08] J. Lott – "Some geometric calculations on Wasserstein space", *Comm. Math. Phys.* **277** (2008), no. 2, p. 423–437.

[LV09] J. Lott & C. Villani – "Ricci curvature for metric-measure spaces via optimal transport", *Ann. of Math. (2)* **169** (2009), no. 3, p. 903–991.

[McC97] R. J. McCann – "A convexity principle for interacting gases", *Adv. Math.* **128** (1997), no. 1, p. 153–179.

[Sch70] L. Schwartz – "Systèmes projectifs de mesures et théorème de Prokhorov", Séminaire Laurent Schwartz 1969–1970: Applications radonifiantes, Exp. No. 1, Centre de Math., École Polytech., Paris, 1970, p. 5.

[ST09] W. Schachermayer & J. Teichmann – "Characterization of optimal transport plans for the Monge-Kantorovich problem", *Proc. Amer. Math. Soc.* **137** (2009), no. 2, p. 519–529.

[Stu06] K.-T. Sturm – "On the geometry of metric measure spaces. I, II", *Acta Math.* **196** (2006), no. 1, p. 65–131, 133–177.

[TY09] A. Takatsu & T. Yokota – "Cone structure of L^2-Wasserstein spaces", arXiv:0812.2752, 2009.

[Vil03] C. Villani – *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.

[Vil09] ———, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009, Old and new.

Jérôme Bertrand, Institut de Mathématiques, Université Paul Sabatier, 118 route de Narbonne, F31062 Cedex 9 Toulouse, France
E-mail : `bertrand@math.univ-toulouse.fr`

Benoît Kloeckner, Université de Grenoble I, Institut Fourier, CNRS UMR 5582, BP 74, 38402 Saint Martin d'Hères cedex, France
E-mail : `benoit.kloeckner@ujf-grenoble.fr`